I. MODIFICATIONS OF THE RADIAL DISTRIBUTION METHOD OF INTERPRETATION OF ELECTRON DIFFRACTION PHOTOGRAPHS OF GAS MOLECULES

II. THE ELECTRON DIFFRACTION INVESTIGATION OF THE STRUCTURE OF BENZENE, PYRIDINE, PYRAZINE, BUTADIENE, CYCLOPENTADIENE, FURAN, PYRROLE, AND THIOPHENE

Thesis by

Verner Schomaker

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1938

TABLE OF CONTENTS

SECTION I

MODIFICATION OF THE RADIAL DISTRIBUTION METHOD OF INTERPRETATION OF ELECTRON DIFFRACTION PHOTOGRAPHS OF GAS MOLECULES

A. Introduction

The problem of the interpretation of data obtained from experiments on the diffraction of electrons by gas molecules[1], for the purpose of obtaining information about interatomic distances and bond angles in a large number of compounds, must be attacked in a way that is both efficient and reasonably accurate. Two such ways exist. The first, called the "visual" method[2], consists of the qualitative and quantitative comparison of the diffraction pattern, as determined by visual inspection of the photographic record, with that calculated with the aid of a theoretical intensity function for an assumed molecular model. The second, the "radial distribution" method[3], employs the same data as the first, but by an inversion of the theoretical formula leads directly to the desired information. Its advantages of simplicity and economy of effort have, however, been generally overbalanced by the greater power and reliability of the visual method.

1. For a comprehensive review on the investigation of structure of gas molecules by the electron diffraction method, see L.O. Brockway, Rev. Mod. Phys., 8, 231, (1936).

2. R. Wierl, Ann. d. Physik, 8, 231, (1936).
L. Pauling and L.O. Brockway, J. Chem. Phys., 2, 867, (1934).
For relatively unimportant and little used variants of the method see Brockway, loc. cit.

3. L. Pauling and L.O. Brockway, J.A.C.S., 57, 2684, (1935).

In this section modifications, involving v y little additional labor, are presented, which, tog er with the generally more complete data available at resent, greatly increase the efficacy of the radial di. ibution method, and render it equal to the visual m hod in importance as a tool for the treatment of elec n diffraction data in the study of molecular structur With these modifications, it can generally be depende upon to give interatomic distances which suggest a model that needs only to be refined somewhat by the application of the visual method.

The greater part of this section is devoted to a discussion of the theory, developed chiefly by Dr. Degard, which is found to underlie these modifications, and of an effort to estimate the importance of the approximations involved in the radial distribution method as it now stands. The common errors of the older radial distribution method can be partially understood with the help of the conclusions based on this discussion, and further modifications of the radial distribution method are suggested.

I am greatly indebted to Dr. Charles Degard for his several very essential contributions to the work

of this section, and to Dr. Simon Bauer for helpful early discussions, and I owe thanks to Professor Pauling, especially, for constant, and frequently necessary, encouragement, as well as for valuable help with the work itself.

B. General Theory

1. Definition of $D(\ell)$; Derivation of the general radial distribution formula.

For the purposes of this section it is desirable to generalize somewhat the usual expression[1] for the intensity of the cone of electrons elastically scattered at an angle φ with respect to an homogeneous incident beam of electrons of wave length λ by randomly oriented gas molecules. This is done by introducing a formally continuous distribution function $D_{ij}(\ell)$, which is proportional to the probability of finding in the molecule and atom i and an atom j a distance ℓ apart. The result is:

(1) $$I_{molecular}^{(s)} = k' \sum_{\substack{i,j \\ i>j}} \int_0^\infty d\ell \, \ell^2 D_{ij}'(\ell) \frac{(Z_i - F_i)(Z_j - F_j)}{Z_i Z_j} e^{-s^2} \frac{\sin s\ell}{s\ell}$$, where

k' is the uninteresting constant $2 I_0 \left(\frac{8\pi^2 m e^2}{h^2}\right)^2$, Z_i is the atomic number of the i-th atom and F_i , its X-ray form factor, and $s = \frac{4\pi \sin \vartheta/2}{\lambda}$. $D_{ij}(\ell)$ is normalized in the following manner.

(2) $$\int_0^\infty \ell^2 D_{ij}'(\ell) \, d\ell = \int_0^\infty D_{ij}(\ell) \, d\ell = Z_i Z_j$$

1. See, for example, Brockway, Rev. Mod. Phys., 8, 237, (1936).

It is clear that (1) automatically provides for the treatment of rigid, as well as non-rigid, molecules. Equation (1) may be rewritten as

(3) $$I(s) = K' \int_0^\infty dl\, l^2 \frac{\sin sl}{sl} \sum_{i>j} \underbrace{\{ D_{ij}(l) \}}_{\text{fcn. of } l} \underbrace{\left\{ \frac{(Z_i - F_i)(Z_j - F_j)}{Z_i Z_j} \right\}_s}_{\text{fcn. of } s}$$

, but this is not enough for our purpose; it is necessary to write the integrand of I(s) as a product of the form, $\sin(sl)_{u(l)}\, v(s)$. It is convenient to introduce the following definitions:

(4) $$D(l) = \sum_{i>j} D_{ij}(l)$$

(5) $$\sum_{i>j} D_{ij}(l) \frac{(Z_i - F_i)(Z_j - F_j)}{Z_i Z_j} = D(l) \left(\frac{Z - F}{Z} \right)^2 = D(l) f^2(s,l)$$

Evidently the desired separation cannot in general be accomplished, even if the structure of the molecule under consideration is known; and obviously, of course, it is impossible if nothing is known about the structure of the molecule at all. However, it is seen on examination of the form factors, F_i, that $\frac{Z_i - F_i}{Z_i}$ is nearly the same function of s for all atoms[1], so that for any set of atoms it is possible to choose a function $f(s)$ which will be a reasonably good representation of $f(s,l)$.

1. This fact is used in the visual method of interpretation of electron diffraction photographs.

Usually, indeed it will be found satisfactory to ignore the variability of f (s). As the result of all this, (3) may now be written

$$(6) \qquad I(s) = K' \int_0^\infty \frac{f^2(s)}{s^5} \, l \, D'(l) \sin sl \, dl$$

It is justifiable, as will be seen later, to allow s to have all positive values instead of imposing the functional limit which arises from its definition. This being true, (6) satisfies the requirements of a Fourier integral and can be inverted in the usual way. For if[1]

$$g(u) = \sqrt{\frac{2}{\pi}} \int_0^\infty f(t) \sin ut \, dt \qquad \text{, where } f(t)$$

satisfies certain requirements, and u, as well as t, may have all positive values, then,

$$f(t) = \sqrt{\frac{2}{\pi}} \int_0^\infty g(u) \sin ut \, du$$

The application of these formulae to (6) yields the desired result[2],

$$(7) \qquad D'(l) = K \int_0^\infty \frac{s^6 I(s)}{f^2(s)} \frac{\sin sl}{sl} \, ds \qquad \text{, where } K = \frac{2}{\pi K'} .$$

1. Courant and Hilbert, "Methoden der Mathematischen Physik," Vol. I, Julius Springer, Berlin, page 68.
2. This is a slight extension of Degard's formula[3] (6.3)
$$D(l) = K \int_0^\infty \frac{s^6 I(s)}{(Z-F)^2} \frac{\sin sl}{sl} \, ds \; ; \; \int_0^\infty l^2 D(l) \, dl = 1$$
which is applicable to homoatomic molecules only, and is probably more generally useful than the kind of treatment of (6.3) given by Degard in the case of carbon tetrachloride.
3. Thesis, University of Liége, 1937.

2. Substitution of a summation over discrete terms for the integration over all values of s.

In the radial distribution method of Pauling and Brockway[1] equation (7) is replaced by a sum, of which each term represents the part of the integral over a corresponding feature of the observed intensity function. Although this procedure was found to be satisfactory, its use in calculating radial distribution functions with visually estimated intensities did not make clear the significance of the approximation involved. This, the later work of Degard has done.

Degard carefully determined the relative molecular intensities as a function of s, and evaluated the integral of the radial distribution function numerically[2]. He then showed that quite precisely the same result is given by the following sum, the terms of which are the same as those of the older formula except for the exponential factor.

$$(8)\quad D'(l) = \sum_{k=1}^{\infty} C_k \, e^{-b_k l^2} \frac{\sin s_k l}{s_k l}$$

1. J.A.C.S., 57, 2684, (1935).
2. Thesis, University of Liége, 1937.

Here the k-th term closely approximates the part of the integral between the sub-limits s_k' and s_k'', the zeros of the molecular intensity function adjacent to the maximum, or minimum, s_k, and the values of C_k and b_k are determined by the height and width of the feature in question. In order to discuss the theoretical significance of this result, and to discover its generalization for more complicated molecules, Degard derived the mathematical expression for the intensity function which exactly corresponds to the radial distribution function (8). In the next paragraph Dr. Degard's treatment is reproduced--with some simplification of the mathematical discussion, and other slight changes.

The theoretical intensity function which is related to (8) is found by substituting for $D'(l)$ in (7) its expression by equation (8):

$$(9)\quad I'(s) = \frac{4f^2(s)}{s^5}\int_0^\infty dl\, l \sin sl \left\{\sum_k C_k e^{-b_k l^2} \frac{\sin s_k l}{s_k l}\right\} = \sum_k I_k'(s)$$

where

$$(10)\quad I_k'(s) = \mathcal{K}\int_0^\infty e^{-b_k l^2} \sin sl \sin s_k l\, dl\,; \quad \mathcal{K} = \frac{4f^2(s)\, C_k}{s_k s^5}$$

$$(11)\quad = \mathcal{K}/2 \int_0^\infty e^{-b_k l^2} \left\{\cos(s-s_k)l - \cos(s+s_k)l\right\} dl$$

By the well known integral,

$$\int_0^\infty e^{-ax^2} \cos bx\, dx = \frac{\sqrt{\pi}}{2\sqrt{a}} e^{-b^2/4a} \quad ; \sqrt{a} > 0$$

, this becomes

$$(12) \quad I_K'(s) = \frac{\sqrt{\pi}\, \mathcal{K}}{4 b_K} \left\{ e^{-\frac{(s-s_K)^2}{4 b_K}} - e^{-\frac{(s+s_K)^2}{4 b_K}} \right\}$$

The second term is negligible for the values of s_K and s which are of interest. Therefore,

$$(13) \quad I_K'(s) = \frac{K\sqrt{\pi}\, f^2(s)\, C_K}{4\sqrt{b_K}\, s^5 s_K} e^{-\frac{(s-s_K)^2}{4 b_K}}$$

On multiplying (13) by $\frac{s^5}{K f^2(s)}$ to obtain

$$(14) \quad I_K^0(s) = \frac{s^5}{K f^2(s)} I_K'(s) = \frac{\sqrt{\pi}\, C_K}{4\sqrt{b_K} s_K} e^{-\frac{(s-s_K)^2}{4 b_K}}$$

, it is seen that the distribution of intensity corresponding to (8) is the same as the theoretical intensity function I(s) insofar as the <u>reduced</u> theoretical intensity function

$$(15) \quad I^0(s) = \frac{s^5}{K f^2(s)} I(s) = \int_0^\infty dl\, l\, O(l) \sin sl$$

can be represented as a sum of Gaussian distributions:

$$(16) \quad I^0(s) = \sum_{K=1}^{\infty} C_K' e^{-\frac{(s-s_K)^2}{4 b_K}} \quad ; \quad C_K' = \frac{1}{4}\sqrt{\frac{\pi}{b_K}} \frac{C_K}{s_K}$$

It will be shown that this can be done quite well by letting the s_K be the s values of the maxima and minima of the reduced theoretical curve, and choosing the b_K and C_K' to correspond to their widths and heights.

Degard used a term for each feature (maximum or minimum) of the reduced theoretical curve, and, moreover, assumed that a peak, say, is best represented if C'_κ is proportional to the height of the peak, and $b_\kappa = 0.361/(s_\kappa' - s_\kappa)^2$, that is to say, if the Gaussian and the peak have the same half-width $(s'_\kappa - s_\kappa)$. Figure Ia shows this representation of the specially simple reduced theoretical function $_\lambda(\sin s\ell)$ for a rigid diatomic molecule, where the C'_κ have been given the value 1.03. It is seen that the theoretical curve (dashed) and the sum of the Gaussian representations (heavy, solid curve) are indeed very nearly alike. Figure Ic shows the representation which is obtained by using Gaussian terms for the ridges only, in the following manner. It is required that when the sum of the Gaussians coincides with the maxima and minima of the reduced theoretical curve (after subtracting a constant, non-essential part), the representation shall as closely as possible approach the theoretical curve at its points of inflection, which, if we now regard it as a series of peaks, are the points of half-maximal ordinates. If we write: $I^b(s) = \sum_\kappa c_\kappa' e^{-\frac{(s-s_\kappa)^2}{4b_\kappa}}$,

these conditions may be expressed so:[1]

$$c_k'(1-2e^{-\frac{4(s_k'-s_k)^2}{4b_k}})=2 \quad ; \quad c_k'(e^{-\frac{(s_k'-s_k)^2}{4b_k}}-2e^{-\frac{4(s_k'-s_k)^2}{4b_k}}) \approx 1,$$

and their best solution gives

$$c_k' = 0 \quad ; \quad b_k = 0.54(s_k'-s_k)^2.$$

The representation fails in that it has a too small ordinate at points corresponding to the inflections of the theoretical curve. Choices of b_k from $0.45(s_k'-s_k)^2$ to $0.60(s_k'-s_k)^2$ are very nearly (within 2%) as good as this optimum. For representations with both maxima and minima it seems reasonable to demand that when the sum of the Gaussian terms fits the reduced theoretical curve at its maxima, minima, and nodes, it should in addition fit as closely as possible at the half-maximal points of the ridges and valleys. The problem of finding the best value of b_k turns out to be exactly the same as before, and the result is formally the same; one must only place the required interpretation on the term "half-width." The representation obtained with this choice of b_k is illustrated in Figure Ib, where it should be noticed that a slight change of scale would serve to show that this choice of b_k does give

1. In formulating these conditions it has been assumed that the Gaussian has a negligible value when its argument is equal to four times its half-width. Although this assumption is not precisely true, it has no appreciable effect on the determination of b_k. The assumption has not been made in plotting the figures.

a better representation than that of Degard, although the difference is probably not significant.

The practical problem of representing the reduced intensity function of a polyatomic molecule is more difficult to discuss. It will be sufficient to remark that the very good representation obtainable for the simple sine function indicates that the sum of Gaussian terms is probably suitable for representing the more complicated Fourier series also. Probably it is best always to represent both maxima and minima of the Fourier series in the hope of obtaining from the sum of two fairly suitable representations one that is somewhat better than either alone. Moreover, it seems that relatively small values of b_κ should be taken, when the representation of both maxima and minima is used, in order to reduce the overlapping of adjacent Gaussian terms, and consequently also the dependence on their nice cancellation. For in the more complicated cases it cannot be expected that this cancellation will be as satisfactory as it is in the example discussed above. Asymmetric peaks, or peaks with "shelves" are specially difficult. Generally a term with a small coefficient can be taken to represent the asymmetry if it is great; and terms corresponding to both the main

peak, and to the shelf can be taken to represent a composite feature of the intensity function.

In this section it has been proved that the general radial distribution function can be replaced by a sum

(8) $D'(\ell) = \sum_{k=1}^{\infty} C_k e^{-b_k \ell^2} \frac{\sin s_k \ell}{s_k \ell}$, which is

equivalent to it insofar as the <u>reduced</u> intensity function can be represented by a sum of Gaussian terms

(16) $I'_0(s) = \sum_{k=1}^{\infty} C'_k e^{-\frac{(s_k - s_k)^2}{4 b_k}}$; $C'_k = \frac{1}{4}\sqrt{\frac{\pi}{b_k}} \frac{C_k}{s_k}$.

According to equation (16) the coefficient C_k of a term of the radial distribution function should be proportional to the height of the corresponding term of the representation of the reduced intensity function, to s_k and to the square root of b_k . It may be inferred from the detailed treatment of the example of a rigid diatomic molecule that the best value of b_k is given by $b_k \geq 0.54 (s'_k - s_k)^2$ where $(s'_k - s_k)$ is the half-width of the k-th feature of the reduced intensity curve.

Finally, it can be seen from the form of the reduced intensity function (15) that for a rigid molecule (the distribution $D(\ell)$ is discrete) it has no secular dependence on s . From this statement

and equations (8) and (16) follows the important corollary that on the <u>average</u> the terms of the radial distribution function (8) for a rigid molecule are <u>equally</u> important.

3. The "artificial temperature factor"; its relation to the use of an only finite number of terms in the radial distribution summation.

Application of the results so far obtained to observed diffraction data does not yield thoroughly satisfactory radial distribution functions. For example, Degard found that even when carefully determined intensities are used, the radial distribution function calculated for carbon tetrachloride in this way shows many relatively weak, false maxima in addition to those corresponding to the C-Cl and Cl-Cl distances in the molecule. Dr. Degard found that these spurious features can be eliminated by using in the radial distribution function instead of the true intensities I(s), quantities $I(s)\, e^{-as^2}$ in which a is chosen so that $e^{-as_m^2} = 1/10$ where s_m is the upper limit of the integration or summation. The form of the expression e^{-as^2} suggests the name "artificial temperature factor," inasmuch as the intensity function for a vibrating molecule differs from that for a rigid molecule

by just such a factor. The artificial temperature factor rapidly and progressively decreases the importance of the part of the distribution function corresponding to large values of s; its action is therefore that of effecting a smooth compromise between the theoretical demand that parts of the radial distribution arising from all values of s should on the average be equally important, and the practical necessity of giving s a finite upper limit corresponding to the greatest observed angle of scattering.[1]

The effect of the artificial temperature factor on the significant peaks of the radial distribution function can be anticipated from the circumstance mentioned in the last paragraph--that this factor is of the same form as the factor which occurs in the theoretical intensity function for a vibrating molecule. A distribution function involving the artificial temperature factor corresponds to a model in which every interatomic distance is subject to the actual mean square variation in the molecule, plus that corresponding to the artificial temperature factor, i.e., $\overline{\delta l_{ij}^2}' = \overline{\delta l_{ij}^2} + 2a^2$. Consequently, if it is desired

1. When this is done, the terms corresponding to large s values become unimportant, and the extension of the upper limit of variability of s to infinity, made in the derivation of the radial distribution formula, is justified.

2. See Brockway, Rev. Mod. Phys., 8, 238, (1936) for a brief discussion of the temperature factor.

to determine the amplitude of the molecular vibrations from the width of a peak of the radial distribution function, it is necessary to subtract a term corresponding to the "artificial amplitude" introduced by the artificial temperature factor. This consideration is valid if the artificial temperature factor has been chosen large enough so that the real and artificial temperature factors together render unimportant the parts of the radial distribution function corresponding to $s > s_m$, which have been neglected, and if the peak representing the distance in question is well separated from its neighbors.

Degard has given a different, more detailed treatment[1] of this question based on an approximate evaluation of the incomplete radial distribution function -

(17) $$D'(l) = K \int_0^{s_m} \frac{s^6 I(s)}{f^2(s)} \frac{\sin sl}{sl} \, ds$$ in which

I(s) has been given the form

(18) $$I(s) = K' \frac{f^2(s)}{s^4} e^{-as^2} \frac{\sin sl'}{sl'}$$ applicable

to a non-rigid diatomic molecule, and suitable when an artificial temperature factor is used. Degard's integration is valid only for $(l - l') \approx 0$, so that it can only

1. Thesis, Chapter 8. Note that in this paragraph b does not have the same meaning as elsewhere in this thesis.

give information about the near-maximal region of the main peak of $D'(\ell)$. Degard has calculated the half width of this peak at <u>three-quarter-height</u> as a function of a and s_m, and has plotted the quantity b of a corresponding Gaussian representation ($e^{-\frac{(\ell-\ell_0)^2}{4b}}$) of this peak, as a function of a and s_m. b is large and very nearly independent of a for small values of s_m; for larger values of s_m, b decreases rapidly, and asymtotically approaches the value a. And the greater the value of a, the smaller is the value of s_m for which $b \approx a$. Unfortunately Degard's curves seem to be subject to error, for certain theoretical relations among them are not precisely satisfied.

For the case when $s_m \to \infty$ the value of the integral is exactly known and the resulting function[1] is of importance here, and is of great interest in connection with the theory of the temperature effect inasmuch as it gives precisely the error incurred in omitting the higher order terms of that theory, and suggests a correction which is almost exact. However, this correction will not be discussed in this thesis.

The substitution of (18) in (17) gives

(19) $$D'(\ell) = \frac{KK'}{\pi\ell_0}\int_0^{s_m} e^{-as^2} \sin s\ell_0 \sin s\ell \, ds$$

1. Degard, Thesis, Chapter 7.

which, when $s_m \to \infty$, is found to be

(20) $$D'(\ell) = \frac{4\pi' \sqrt{\pi}}{\ell\ell_0 2\sqrt{a}} \left\{ e^{\frac{-(\ell-\ell_0)^2}{4a}} - e^{\frac{-(\ell+\ell_0)^2}{4a}} \right\}$$

in which the second term is negligible for values of and a ordinarily encountered. Consequently,

(21) $$D(\ell) = \ell^2 D'(\ell) = \frac{1}{\sqrt{\pi a}} \frac{\ell}{\ell_0} e^{\frac{-(\ell-\ell_0)^2}{4a}}$$ is the

distribution of scattering matter corresponding to (18), with essentially the norm ($\int_0^\infty D(\ell)d\ell = 1$) and the half-width ($\frac{1}{2}\sqrt{\delta\ell_{ij}^2} = a$) assumed in formulating (18), the factor ℓ/ℓ_0 being unimportant in these respects. This factor is effective, however, in that it shifts the peak of the radial distribution function so that $\ell_{max}/\ell_0 > 1$: the artificial temperature factor gives rise to error, which, however, may be eliminated by taking, as the interatomic distances given by the radial distribution method, the peaks of $\ell D'(\ell) = D(\ell)/\ell$ rather than those of $D(\ell)$. The real temperature factor corresponding to the actual intensities, differs from its approximate form e^{-as^2} in such a way of course, that this shift does not occur, and therefore the above conclusions must not deceive us into mistrusting the faithfulness of the theoretical radial distribution function.

It seemed desirable for the present thesis to supplement Degard's discussion of the effect of a finite s_m, and of the temperature factor, with the following further considerations. It is impossible, as has already been indicated, to find a simple expression of the "incomplete" radial distribution function for the case when I(s) includes a temperature factor. However, for our purposes, $\frac{1+\cos cs}{2}$ can be made to be a sufficiently good approximation of e^{-as^2}, as is illustrated by Figure II. We shall investigate the nature of the radial distribution function for a diatomic molecule when c is given various values. For the comparison of the results of this investigation with the corresponding radial distribution functions involving the correct form of the temperature factor, it is necessary to find what value of a makes e^{-as^2} and $\frac{1+\cos cs}{2}$ most nearly the same for $0 \leq s \leq s_m$. In Figure II where $a = 0.28\, c^2$, ($b = 0.01$; $c = 0.189$), the two forms of the temperature factor have been made to have the value $\frac{1}{2}$ for the same value of s. This relation of c to a is seen to give a good representation of the true temperature factor over the range $1 \geq e^{-as^2} \geq 0.25$. For the range $1 \geq e^{-as^2} \geq 0.1$, a better average representation is obtained if $a = 0.30\, c^2$.

And if it is desired to obtain a representation over a still larger range of the temperature factor, a must be given a somewhat larger value still.[1]

The theoretical radial distribution function for a diatomic molecule having the temperature factor discussed in the last paragraph can be set up and evaluated straightforwardly. We have

(22) $I(s) = \frac{\kappa' f^2(s)}{s^4}\left(\frac{1+\cos^2 cs}{2}\right)\frac{\sin l's}{l's}$, and hence,

from (7):

$$D'(l) = k\int_0^{s_m} \frac{(1+\cos^2 cs)\sin l's \sin ls}{ll'}\,ds \; ; \; k = \frac{kk'}{2} = \frac{1}{\pi}$$

(23)
$$= \frac{k}{2ll'}\int_0^{s_m} \{1+\cos^2 cs\}\{\cos(l-l')s - \cos(l+l')s\}\,ds$$
$$= \frac{k}{2ll'}\int_0^{s_m} \{\cos(l-l')s - \cos(l+l')s + \tfrac{1}{2}[\cos(l-l'+c)s + \cos(l-l'-c)s - \cos(l+l'+c)s - \cos(l+l'-c)s]\}\,ds$$

In order that this function shall correspond to integration over a definite number of maxima and minima of the reduced intensity function, we set

(24) $s_m = \frac{n\pi}{l'}$. Then,

(25) $$D'(l) = \frac{k}{4ll'}\left[\frac{4(-1)^n \sin n\pi\frac{l}{l'}}{l'\{(\frac{l}{l'})^2-1\}} + \frac{\sin\frac{n\pi}{l'}(l-l'+c)s}{(l-l'+c)} + \right.$$
$$\left. + \frac{\sin\frac{n\pi}{l'}(l-l'-c)}{(l-l'-c)} - \frac{\sin\frac{n\pi}{l'}(l+l'+c)}{(l+l'+c)} - \frac{\sin\frac{n\pi}{l'}(l+l'-c)}{(l+l'-c)}\right.$$

1. Although the representation in this case is no longer very accurate, the discrepancy is probably unimportant for this discussion.

On making the following substitutions

(26) $\frac{\ell}{\ell'} = L$; $\frac{c}{\ell'} = C$, and simplifying,

(27) $$D'(\ell) = \frac{(-1)^n k}{\ell'^3 L}\left\{\frac{\sin n\pi L}{L^2-1} + \frac{(L^2+C^2-1)\sin n\pi L \cos n\pi C - 2LC \sin n\pi C \cos n\pi L}{(L^2+C^2-1)^2 - 4L^2C^2}\right\}$$

is finally obtained as the exact expression of $D'(\ell)$ for a diatomic molecule and the approximate temperature factor $\frac{1+\cos cs}{2}$. This result represents n features (maxima and minima being counted separately) of the reduced intensity curve. In the absence of a temperature factor ($c=0$), (27) becomes much simpler:

(28) $$D'(\ell) = \frac{2(-1)^n k}{\ell'^3 L} \cdot \frac{\sin n\pi L}{L^2-1}$$

$D(\ell) = \ell^2 D'(\ell)$ has been evaluated for $n=5$, 10, and 20 and $C=0$, $1/n$, and, for $n=10$, $0.8/n$. The results are plotted to an arbitrary scale in Figure III where the convenient notation $D_y^x(\ell)$ has been adopted to represent $D(\ell)$ for $n=x$ and $C=y$. It should be noted that $C=0$ corresponds to a unit temperature factor; $C=0.8/n$, to a factor equal to 0.1 when $s=s_m$; and $C=1/n$, to an approximate factor equal to zero when $s=s_m$. The last case is the extreme limit of the usefulness of the approximation of e^{-as^2} by $\frac{1+\cos cs}{2}$, and really best corresponds to a temperature factor

which has only a very small value (say 0.03) when $s = s_m$. A part of the peak of $D'(\ell)$, and of $\ell D'(\ell)$, for each choice of n and C is also shown in Figure III. Since the height of the main peak of $D(\ell)$ is proportional to n,[1] it is convenient to normalize by dividing by n, and this has been done.

These theoretical curves are of interest in connection with radial distribution functions in general, because for any molecule the distribution function may be regarded as a superposition of several curves of this type. This radial distribution function will be satisfactory if each of the component parts has the form of a peak with level plains at either side, and if the various peaks are not too closely spaced to be well resolved. And this will be true also in the case where the radial distribution function is calculated as the sum of several Fourier terms, with suitable coefficients, if the conditions necessary for this approximation, as discussed in section B2, have been satisfied.

The features of the curves, of Figure III, which appear to be specially important will now be discussed. Those obtained without a temperature factor are seen to

1. This is easily seen, from (28) for instance, by evaluating the indeterminant expression $D'(L=1)$ by L' Hospital's rule.

have a central important peak and neighboring lesser peaks which are progressively smaller, the further they are removed from the main peak. The nodes are equally spaced except at the main peak, which, at its nodes, is just twice as wide as the other features. The functions $D^{n}_{1/n}(\ell)$ have main peaks only half as high as those for which C 0, while the remainder of the curve follows the L axis very closely, the corresponding minor features of the first curves having almost entirely disappeared. The function $D^{n}_{0.3/n}(\ell)$ has a somewhat higher maximum value, while the spurious features have become negligible for it also[1]. This is the desired effect, for the superfluous features of the radial distribution function in the practical case, when no temperature factor is employed, are very confusing and render its interpretation very difficult. It may be well to remark that probably any other function which suitably reduces to very small importance parts of the radial distribution sum or integral corresponding to the outer rings of the diffraction photograph, would serve the same purpose as the artificial temperature factor. It may also be worthwile to mention again that the artificial and real temperature factors act in exactly

1. Perhaps a less drastic factor still would serve. However, this one, chosen to have the value 1/10 for $s - s_m$, according to the suggestion of Degard, is seen to be satisfactory, and is perhaps as satisfactory as any for general use.

the same way, so that as the real temperature factor becomes more important, the artificial temperature factor may become less so.

The width of the peak of the radial distribution function is of interest, for it determines the resolution which it is possible to obtain when several of the simple curves (for different l''s) are added together in the practical case. The half-widths of the peaks of the curves of Figure III are found to be 0.6/n for C = 0; 1.0/n $\cdot l'$ for C = 1/n; and, presumably, 0.8/n $\cdot l'$ for C 0.8/n.[1] That these results for the cases C = 1/n and C = 0.8/n are in agreement with the general statements concerning the width of a radial distribution peak given earlier in this section (page B15), can be seen from the following table which gives the half-width of the peak of the radial distribution curve, the value of a ($a = 0.30\ c^2$) corresponding to C, and the half-width of the distribution of scattering matter corresponding to the temperature factor e^{-as^2} ($\overline{\delta l_{ij}^2} = 2a$; for a Gaussian distribution $\chi_{1/2} = 1.176\sqrt{\overline{\chi^2}}$).

1. The generalization for C = 0.8/n is justified by the apparent circumstance that the various curves for a given value of C (e.g. 1/n) are the same with regard to general appearance and width of the peak except for the scale of (L-1). That is, they are similar functions of n(L-1). That this should be true is not obvious from the form of (25).

Curve	Observed half-width	a	Half-width calculated from a
$D^{5}_{1/5}$	0.19	0.012	0.183
$D^{10}_{1/10}$	0.10	0.003	0.091
$D^{20}_{1/20}$	0.05	0.0075	0.046
$D^{10}_{08/n}$	0.08	0.0019	0.073

For general radial distribution functions the value of n to be used in estimating the expected half-width of a particular peak is the number of maxima and minima of the corresponding component of the reduced intensity function which lie between the origin and s_m. It is convenient to state the above results in another form for practical use. This is easily done by noting that

(29) $n_i = \frac{s_m \ell_i'}{\pi}$, and making use of the definition of C, (26). The results are given in the following table.

Theoretical Half-widths of Peaks of the Incomplete Radial Distribution Function with a Temperature Factor

c	Temperature	Factor (corresponding)	Half-width
0	1	1	$1.9/s_m$
$2.51/s_m$	0.1	0.1	$2.5/s_m$
$3.14/s_m$	0.0	0.03	$3.1/s_m$

Two peaks of equal height may be expected to be well resolved if their separation be greater than three times their average half-width; or, resolved, but shifted toward each other if their separation be about twice their average half width. The predictions of this table are in general agreement with experience for radial distribution curves calculated by the modified method described in section C. This is true with regard to the magnitudes of half-widths as well as their equality for all the well resolved peaks, having like real temperature factors, of any one radial distribution function.

Finally we direct our attention to the positions of the maxima of the curves of Figure III. It is seen that in general the maxima of $D(\ell)$ occur for L slightly greater than the theoretical value of unity, while the maxima of $D'(\ell)$ occur when L is slightly less than one, and those of $\ell D'(\ell)$ are at $L = 1$. That this should be true is evident from the form of (27) for $C = 0$, and by (21), for any choice of the usual temperature factor, also if s_m be large enough. From (27) it is evident that when these conditions are not met there may well be some shift of $\ell D'(\ell)$, but this effect we expect to be small. The antibatic dependence on n of the

magnitudes of the shifts cannot be simply expressed. The values of these shifts as given in the following table are obviously only approximate; better ones could best be obtained by a more precise calculation of (27), since the analytical treatment of (27) is apparently very tedious.

Shifts of Maxima
of Incomplete Radial Distribution Functions

n	C	Shift of Maximum of D (ℓ)	Shift of Maximum of D' (ℓ)
5	0	+0.007	-0.020
10	0	+0.004	-0.004
20	0	-----	-----(very small)
5	1/5	+0.025	-0.033
10	1/10	+0.007	-0.008
20	1/20	+0.002	-0.002
10	0.8/10	+0.007	-0.005

The error depending on this effect can probably be eliminated in general by taking the positions of the maxima of $\ell D'(\ell)$ instead of those of either $D(\ell)$ or $D'(\ell)$. In general, it may be expected that these shifts may vary from about a per cent for $n = 6$, to about a tenth of a per cent for $n = 20$, and to have the sign found in the above examples.

B. Applications

1. The method of Pauling and Brockway; the transition to a modified method.

The radial distribution formula as derived by Pauling and Brockway[1] was simplified by them to the following form for practical use as an aid in the determination of molecular structures of gas molecules by electron diffraction.

$$(30)\quad D'(\ell) = \sum_k I_k \frac{\sin s_k \ell}{s_k \ell}$$

Here I_k is the visually estimated intensity of the kth ring of the electron diffraction photograph, and the other symbols are already familiar. This function has been applied in the study of many molecules and has enjoyed success, especially for molecules with very few, and well separated, interatomic distances. When these conditions are met, the function yields curves whose maxima occur at values of ℓ in good agreement with those obtained by the visual method of interpretation. There are slight discrepancies among the interatomic distances found by (30), however, and it will be attempted to discuss these later. But great difficulty is met when the molecule under consideration involves a considerable number of significant interatomic

1. J. A. C. S., 57, 2684, (1935).

distances, for the peaks obtained from (30) are excessively broad, and consequently the "resolving power" is poor.

The work of this paper was instigated by Professor Pauling's suggestion that the radial distribution method could be improved by altering the treatment of the coefficients of the sin $s_k\ell$ terms of the sum of equation (30). Figure IV shows radial distribution functions calculated from the observed positions and visually estimated intensities of the rings of electron diffraction photographs of thiophene by an expression similar to (30) in which the coefficients I_k of (30) have been replaced by $I_k s^{-1/3}$, I_k, $I_k s^{1/3}$, $I_k s$, $I_k s^2$, $I_k s^4$, in A, B, C, D, E, and F, respectively. Inasmuch as the data are reported in part two of this thesis it will be unnecessary to give them here also. The interatomic distances in thiophene, as finally determined jointly by the visual and radial distribution methods, are indicated by arrows the lengths of which are proportional to the number and atomic-number product of the corresponding interatomic distances. It is seen that as increasing weight is given to the outer rings, by increasing the exponent of s_k in the coefficients, the resolution of the function greatly improves.

There appear several significant peaks which are either poorly resolved, or not resolved at all, in the first three curves, and these correspond well with the corresponding interatomic distances. It is extremely significant that the positions of the greater peaks are affected only very slightly by varying the treatment of the coefficients. This is important in that it shows how the necessarily large uncertainties of the trend or scale of visually estimated intensities cannot give rise to great errors in the determination of interatomic distance by the radial distribution method, and how both the older method and the modified method presented here can give good results. It is also to be noticed that in E and F there are many more maxima than in A, B, and C, and that some of these new peaks appear to have no physical significance.

A similar set of radial distribution functions was calculated from the data for benzene with exactly similar results. The peak corresponding to the distance between carbon atoms at opposite sides of the benzene ring appears at the correct place although it was unfortunately true that it did not appear at all in the distribution calculated by the older method.

2. The Modified Visual Radial Distribution Method.

Comparison of visual estimates of intensities, for various molecules, made by workers in this laboratory, with the correct simplified theoretical curves for these molecules, indicates that the visual intensity is approximately proportional to (1/s) times the "simplified" theoretical intensity, or (s) times the theoretical intensity. It was therefore suggested that the visual radial distribution function should be given the form

(31) $$D'(\ell) = \sum_{k=1}^{k_{max.}} I_{k(vis.)} s_k^3 e^{-a_k s_k^2} \frac{\sin s_k \ell}{s_k \ell}$$ where I(s)

in (7) has been replaced by $I_{k(vis.)}(s)/s^3$, the factor $\frac{1}{f^2(s)}$ has been left out, the integral has been replaced by a summation as in the older method, Degard's artificial temperature factor, in which it is customary to choose a such that $e^{-as_m^2} = 1/10$, has been inserted, and the summation extends over all the peaks of the observed intensity curve except that at s = 0, the terms corresponding to the first two features of the reduced theoretical curve being thus omitted.

Curves calculated by equation (31) are found to retain the good resolution found for E and F of Figure IV,

and in the similar calculations for benzene. Moreover the prominence of the spurious maxima characteristic of E and F is in general greatly reduced as a result of the action of the artificial temperature factor. In part two of this thesis several examples of (31) are given. Through the efforts of friends in this laboratory, for whose special assistance in this regard the author is most grateful, the new function has been tested for many other molecules of diverse types, and has been found to be powerful and generally reliable. It has become especially evident that even when the function is not completely satisfactory, in that the interatomic distances obtained are not entirely compatible with a reasonable molecular model, it is nevertheless true that this set of interatomic distances, together with suitable weight factors, does often lead to a simplified theoretical curve in complete agreement with the appearance of the original data, so that the radial distribution method leads to a solution of the problem which, if not the correct solution, is at least indistinguishable from it on the grounds of the electron diffraction data alone. There persists some difficulty about spurious maxima. These often can be recognized because they appear between two important peaks, and

although perhaps relatively prominent to the eye, have a maximum value of $D'(\ell)$ of approximately zero. The negative values of $D'(\ell)$ probably arise from the omission of the first two terms of the function (and are hence unimportant), while spurious peaks between these minima are undoubtedly due to imperfections of the function. Further discussion of the modified method is given in section C3d.

3. The modified radial distribution method discussed in the last section is related to the general theoretical radial distribution function by a series of approximations, some of which have been justified in the theory given in part B, while others must now be discussed specially. To the first class belong the substitution of summation for integration, and the use of an only finite number of terms. To the second class belong the omission of the factor $\frac{1}{f^2(s)}$, the omission of the part of the integral or summation corresponding to the first terms of the reduced intensity function, the omission of the factors $e^{-b_k \ell^2}$ demanded by the theory for the substitution of a summation for integration, the interpretation of $D'(\ell)$, rather than $D(\ell)$, as giving the importance and values of the interatomic distances, and the effect of these things and of the artificial temperature factor on the positions of the peaks of the radial distribution function. These are

discussed in the following paragraphs.

a. The effect of omitting the first terms of the summation; the factor $\frac{1}{f^2(s)}$.

It should first be noted that in the treatment of actual intensity data, the effect of neglecting the first features of the reduced theoretical curve is nearly equivalent to the omission of $\frac{1}{f^2(s)}$. For, as is illustrated by figure V, this factor is nearly constant except for small values of s, for which it becomes very large, approaching infinity as s approaches zero. Consequently the first terms of the radial distribution function, which, with the factor, are severally of approximately the same importance as each of the other terms, become without it entirely negligible. These parts of the radial distribution function are essentially positive for the range of values of ℓ which are of interest, so that a function from which they have been omitted will be depressed in this region and will show negative values of $D'(\ell)$. Professor Pauling has pointed out that such a radial distribution function, in which the atom form factors have not been considered, is a representation of the actual distribution of scattering matter in the molecule, the electrons having a negative scattering power relative to the nuclei. The results

of Degard[1] are in agreement with these general statements. He found that radial distribution functions for carbon tetrachloride and bromine calculated from accurately measured intensities according to the methods outlined above are everywhere essentially positive if the factor is included, but have appreciable minima adjacent to the peaks if it is disregarded. Perhaps the troughs of the modified radial distribution functions are partly due to this situation.

Since the function $\frac{1}{f^2(s)}$ is not really constant beyond $s \backsim 5$, nor is the part of radial distribution function from small values of s independent of $\mathcal{l}$, it is desirable to consider these points in greater detail, even if it is not possible to demonstrate conclusively that radial distribution functions constructed from visually observed data have anything to do with them. While a suggestion regarding the treatment of the lesser variability of $\frac{1}{f^2(s)}$ $(s > 5)$ is made in the last section of this discussion, we cannot expect that the estimates of intensity, as they are now made, can be as consistent with regard to scale, over a large range of s, as would be necessary to make a consideration of $\frac{1}{f^2(s)}$, in this region significant. An attempt to estimate the effect on positions of peaks of omitting the first terms of the radial distribution function is given in C3c.

1. Thesis, Chapter 6.

b. The factor $l^2 e^{-bl^2}$.

The modified radial distribution function described in section C2 must, according to the theory, be multiplied by l^2 and further modified by the inclusion of factors $e^{-b_k l^2}$ in order to obtain an expression which may be expected to have peaks whose areas correspond to the probability(multiplied by the product of the atomic numbers) of finding in the molecule two atoms a distance l apart. In the applications to gas molecules, using visually estimated intensities, both of these factors have always been neglected, and indeed it has always been assumed that $D'(l)$ is the function of physical interest. We shall see, however, that these corrections are significant. It turns out that the physically insignificant peaks of the distribution functions which occur for both relatively small and relatively large values of l are eliminated by these corrections, and do not arise from inherent errors of the radial distribution method. (These peaks are already considerably weakened in the transition from the old to the modified method).

If it is assumed that even when the features of the intensity function are not all of the same half-width, it is nevertheless suitable to take an average value

of b_k obtained from the average width of the features of the intensity function, then all of the exponential functions of ℓ become the same and the corrections discussed in this paragraph become simply a factor by which the modified radial distribution function, $D'(\ell)$, is to be multiplied. This is very desirable, for it greatly simplifies the calculation of the corrected modified function. And certainly the use of an average $e^{-b\ell^2}$ is suitable for the present purpose, but it must be understood that the use of such a factor does not improve the representation of the reduced theoretical intensity function upon which the substitution of a summation for integration depends.

The factor $F(\ell) = \ell^2 e^{-b\ell^2}$ has been plotted in Figure VI as a function of s for a range of values of b. The table in this figure gives the value of b, and of the corresponding average values of half-width, $(s_k' - s_k)$, and interatomic distance, ℓ', for each curve. The two values of ℓ' are obtained from the expressions $\ell' = \frac{\pi}{3(s_k' - s_k)}$ and $\ell' = \frac{\pi}{2(s_k' - s_k)}$ for the respective cases when maxima and minima, and maxima only are used. The "average" reduced intensity function is taken as a simple sine curve, and b is obtained from the expression of section B2, $b = 0.5\ (s_k' - s_k)^2$. Since it is impossible to represent all the curves adequately

on the same scale it is well to note that they are interrelated simply:

$$(32)\quad F_b(\ell) = \left(\frac{\ell_b'}{\ell_{b'}'}\right)^2 F_{b'}\left(\frac{\ell_{b'}'}{\ell_b'}\cdot\ell\right) \quad ; \quad \frac{\ell_b'}{\ell_{b'}'} = \sqrt{\frac{b'}{b}} = \frac{(s_k' - s_k)_{b'}}{(s_k' - s_k)_b}$$

It is easily shown that the value of ℓ for which has a maximum value is related to the parameters b, $(s_k' - s_k)$ and ℓ' by the following expression:

$$(33)\quad \ell_m = \frac{1}{\sqrt{b}} = \frac{1.4}{(s_k' - s_k)} = \begin{cases} 0.9\,\ell' & \text{(maxima only)} \\ 1.35\,\ell' & \text{(maxima and minima).} \end{cases}$$

With the help of Figure VI, and the above auxiliary expressions, it is now easy either to correct the modified radial distribution function before plotting it, or to estimate the relative importance of its peaks if it is plotted without correction. Obviously the uncertainty in the semi-quantitative for that reason.

c. Shifts of peaks of the radial distribution function.

It is evident that the corrections discussed under the last two headings imply that the peaks of $D'(\ell)$ will not in general give the correct interatomic distances, but will be shifted somewhat instead. In this section an attempt is made to estimate the shift quantitatively, and it is found that the shifts are inversely dependent on s_m, or directly dependent on

the widths of the peaks of the radial distribution curve. Generally, the effect is somewhat, to spread apart the peaks of $D'(\ell)$. The predicted shifts are too small to be confirmed, except perhaps qualitatively, for radial distributions calculated by the modified method from reasonably complete data. Because functions calculated by the older method show greater errore, an effort is made to test the predictions on them. The application of the theory to the results of the older method is reasonable, because the older method corresponds, at least roughly, to the use of a very drastic temperature factor. A peak of the theoretical R.D. function for an interatomic distance having a Gaussian distribution about a mean ℓ_i, with half-width ω_i is given by

$$(34)\quad D_i(\ell) = e^{-\frac{(\ell-\ell_i)^2}{1.44\omega^2}}$$

and peaks of "incomplete" radial distribution functions are of this form also, if the temperature factor be sufficiently important. We assume, therefore, that in general it is suitable to represent the peaks of a radial distribution function, $D(\ell)$, by (34), and if a peak has not the ideal form an attempt is made to choose ω_i so that the top of the peak is well represented. If ω_i is reasonably small, the peaks of the modified radial distribution function, $D'_m(\ell)$, will have the same

form and width as those of $D_i(\ell)$, and ω may be determined directly from them.

We now write

$$(35)\quad D'_m(\ell) = \frac{\sum D_i(\ell)}{\ell^p e^{-b\ell^2}} = \frac{\sum e^{-\frac{(\ell-\ell_i)^2}{1.44\omega^2}}}{\ell^p e^{-b\ell^2}}$$

where, according to the definition of $D_m'(\ell)$ and the foregoing discussions, p must equal 2 in order that the various peaks of $D'_m(\ell)$ as defined by (35) shall have the same relative importance as in $D'_m(\ell)$ as actually calculated by (31). On the other hand, according to section B3, p=1, if it is desired that the $D'_m(\ell)$ calculated from (35) and (31) give the same interatomic distances.

Now if the peaks of $D(\ell)$ are well resolved, the position of the maximum of any one can be obtained while the others are ignored. Accordingly, we consider

$$(36)\quad D'_{mi}(\ell) = \frac{e^{-\frac{(\ell-\ell_i)^2}{1.44\omega^2}}}{\ell^p e^{-b\ell^2}}$$ for which

$$(37)\quad \frac{dD'_{mi}(\ell)}{d\ell} = e^{b\ell^2 - \frac{(\ell-\ell_i)^2}{1.44\omega^2}} \left\{2b\ell^{-(p-1)} - 2\frac{(\ell-\ell_i)}{1.44\omega^2}\ell^{-p} - p\ell^{-(p+1)}\right\}$$

(37) has zeros at $\ell \approx 0$, $\ell \approx \ell_i$, $\ell \approx \infty$. The root of interest is contained in the second factor which reduces to

$$(38)\quad (1-1.44\omega^2 b)(\ell-\ell_i)^2 + (1-2.88\omega^2 b)(\ell-\ell_i)\ell_i + 0.72\omega^2 - 1.44\omega^2 b\,\ell_i^2 = 0$$

when multiplied by l^{p+1} and expressed as aquadratic in $(l-l_i)$. Since the roots of this quadratic equation are widely separated in absolute magnitude the approximate solution,

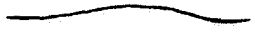
$$(39)\quad (l-l_i)_{max.} \approx \frac{0.72\omega^2(2bl_i^2-p)}{l_i\{1-2.88\omega^2 b\}} \approx \frac{0.72\omega^2}{l_i}\{2bl_i^2-p\}\{1+2.88\omega^2 b\}$$
$$\approx \frac{0.7\omega^2}{l_i}(2bl_i^2-p)$$

is sufficiently accurate. (For cases of interest, the last expression of $(l-l_i)_{max}$ may be in error by 10%.

In the following table are given the calculated errors for the radial distribution functions for a number of molecules calculated by Pauling and Brockway[1] according to their method, together with the observed errors (based on the structures) determined by the visual method. b has been estimated from Figure VI ("maxima only") and a weighted average distance l', while ω is estimated from the radial distribution curves. That it has been necessary to choose p 2 (instead of 1) to obtain the following remarkable agreement is difficult to explain, except as evidence for the very rough character of this discussion. Molecules having only one interatomic distance have been omitted.

1. J. A. C. S., 57, 2684, (1935).

	Molecule	$\overline{\ell'}_{max.}$	b	ω	M - X ℓ_o	M - X $\Delta\ell_{calc.}$	M - X $\Delta\ell_{obs.}$	X - X ω	X - X ℓ_o	X - X $\Delta\ell_{calc.}$	X - X $\Delta\ell_{obs.}$
	CF_4	1.9	.35	.28	1.335	-.031	-.015	.23	2.235	+.025	+.010
	SiF_4	2.0	.30	.26	1.527	-.019	-.017	.25	2.555	+.033	+.030
	CCl_4	2.5	.20	.25	1.74	-.020	-.020	.20	2.87	+.011	-.010
	$SnCl_4$	2.7	.16	.4	1.963	-.044	-.053	.3	3.27	+.027	-.016
	$GeCl_4$	2.5	.20	.3	2.070	-.009	-.033	.3	3.385	+.042	-.055
	$SnCl_4$	2.7	.16	.27	2.295	-.007	+.006	.3	3.76	+.030	+.015
	PF_3	1.8	.40	.22	1.47	-.006	-.050	.22	2.41	+.036	+.040
	PCl_3	2.4	.22	.22	1.98	-.005	-.040	.3	3.08	+.044	-.010
	$AsCl_3$	2.4	.22	.31	2.13	0	-.050	.37	3.43	+.087	+.070
(H_2)	CCl_2	2.3	.24	.24	(1.73)	-.013	-.040	.25	2.935	+.032	+.015
(H)	CCl_3	2.5	.20	.16	(1.77)	-.008	-.010	.31	2.931	+.037	+.001
	Cl_2O	2.1	.27	.27	1.65	-.016	-.006	.29	2.86	+.048	+.004

The part of the radial distribution function which depends on the first features (maximum, or maximum and minimum) of the reduced intensity function is of the form  or  for the two cases mentioned, its maximum height varying from $2/n$ — $4/n$ times the height of main peaks of the radial distribution function, according to the importance of the temperature factor. Omission of this term will lead to a spreading apart of the peaks of much the same nature as that discussed in the last paragraphs. The shifts due to this cause will decrease with increasing s_m for two reasons--the narrowing of the peaks and the decreasing relative importance of the omitted terms. It does not seem worthwhile to give a quantitative discussion of this effect: the last example illustrates the approximate magnitude of this one also, and suggests how either can be discussed.

4. Discussion and Suggested further Modifications.

The study of the radial distribution function for use with visually observed data is not complete. In order to gain a thoroughly satisfactory understanding of this method, and to develop its usefulness toward a maximum, it will be necessary to find out more about the

accuracy of the representation of the reduced intensity function, which is necessary, or obtainable, for the various situations which arise. The efficacy of choosing the appropriate values of $e^{-b_2 s^2}$ for different terms of the summation should be investigated, as is specifically demanded by the theory given by Degard. When this id done the explicit form for the coefficients, given in B2, should be followed, and it would perhaps be well to include in the radial distribution function, terms for the first maximum and minimum of the reduced intensity function. Of cours, good measurements of these features cannot be had but they have to be known only so approximately that this need not cause trouble.

Two suggestions which certainly are of practical importance can be made. First, it has been found useful in the few examples which have been tried, some of which are included in Part II, to include terms for both maxima and minima according to the suggestion of Dr. Simon Bauer. Although it is somewhat difficult to estimate the negative coefficients for these terms, and it is possibly true that visual measurements of "minima" are not as reliable as visual measurements of maxima, these terms do make a reliable and sometimes very valuable contribution to the radial distribution

function, and should therefore be used. Often the spurious peaks of the function from the minima tend to cancel those of the maxima. The great difficulty in the visual estimation of intensities is in establishing their general trend. This difficulty can be avoided by arbitrarily fixing the scale of the "intensity" estimates in such a way that on the average they are constant--they are to represent the amplitude of the reduced theoretical intensity function for a rigid molecule. By thus estimating the coefficients of the reduced intensity function the $\frac{1}{f^2(s)}$ factor is nicely provided for, while the average real temperature factor for the molecule is reduced to unity, so that differences among the real temperature factors corresponding to the several interatomic distances will be left unchanged. With these intensity estimates the coefficients become

(40) $s_k I^o_{k(v's.)} e^{-as_k^2}$, where

$e^{-as_k^2}$ is the usual artificial temperature factor.

It may be expected that the radial distribution method which has been described is as reliable as the visual method (and much more reliable and powerful than the older radial distribution method), except for those

cases in which a radial distribution function from visually observed data must always be inferior. Such are the examples where many of the important features of the intensity function are for some reason unsuited to quantitative treatment. These features can be avoided, as such, in the visual method of interpretation, but in the radial distribution method they must be used equally with the rest of the data. Perhaps the results of the visual radial distribution method must still be regarded with awe and faith, even though the theoretical expressions for the half-width (derived in B3) indicate the expected resolution; the curves of Figure VI help to tell whether a peak at a relatively small or relatively large value of $\mathscr{l}$ is significant, and the work of Degard has provided a clear justification for the intrepid approximations of Pauling and Brockway.

Section II

THE ELECTRON DIFFRACTION INVESTIGATION OF THE STRUCTURE OF BENZENE, PYRIDINE, PYRAZINE, BUTADIENE, CYCLOPENTADIENE, FURAN, PYRROLE, AND THIOPHENE

The determination of values of interatomic distances in molecules has been found to provide much information regarding electronic structure, especially in the case of substances which resonate among two or more valence-bond structures. The interpretation of interatomic distances in terms of the types of bonds involved is made with use of an empirical function formulated originally for single bond-double bond resonance of the carbon-carbon bond (1). There are given in this section the results of the determination of the configurations of two molecules involving Kekule resonance, pyridine and pyrazine, of two molecules with conjugated double bonds, butadiene and cyclopentadiene, and of three five-membered heterocyclic molecules, furan, pyrrole, and thiophene, together with the conclusions regarding their electronic structures reached by application of the bond type - interatomic distance function. In addition, because of the improvement in electron diffraction technique since the study of benzene in these Laboratories four years ago (2) and because of its similarity to pyridine and pyrazine, this substance was re-investigated.

Experimental Method.- The diffraction photographs were prepared with the apparatus and technique described by Brockway (3). Ten or more photographs were made for each substance, the electron wave length used being about 0.0613 Å and the camera distance 10.83 cm. The values of

given in the tables are averages of the values found by visual measurement of ring diameters for ten or more films.

The radial distribution treatment, with the refinement described in the first section of this dissertation, was applied as the first step in the interpretation of the data for each substance. Further selection from among the small range of molecular models compatible with the radial distribution curve for each substance was then made by the usual visual method, involving comparison with simplified theoretical intensity curves for the models under discussion.

The sample of benzene used was prepared by redistilling thiophene-free benzene and rejecting the first and last portions. A similar treatment was given to pyridine (Kahlbaum), pyrrole (Eastman), and thiophene (Eastman). Butadiene was prepared by treating tetrabromobutane in ethyl alcohol solution with granulated zinc and redistilling the product several times. Cyclopentadiene was prepared by distilling dicyclopentadiene; in order that appreciable repolymerization should not take place in the sample the diffraction photographs were made immediately after the distillation (within one hour). Furan was prepared by the decarboxylation of furoic acid and purified by fractional distillation. A sample of pyrazine kindly given us by Professor F. W. Bergstrom was used without further purification.

Benzene, Pyridine, and Pyrazine.- New photographs of benzene were obtained showing four more measurable rings than those used in the earlier investigation in these Laboratories. The radial distribution curve calculated with the q and s values of Table I (Fig. 1) shows three principal peaks, at 1.37 Å, 2.43 Å ($= \sqrt{3} \cdot 1.40$ Å), and 2.83 Å ($= 2 \times 1.415$ Å), representing

the C-C separations for a regular hexagon with edge about 1.39 Å. The small peaks do not provide reliable information about the C-H bond distance. In order to make an approximate evaluation of this distance, intensity curves were calculated for plane hexagonal models with C-C = 1.39 Å and C-H varying between 1.00 and 1.18 Å at intervals of 0.03 Å. These curves (Fig. 2) differ significantly only in the region from the third to the sixth ring. (Note than in the tables a feature of the photographs is for convenience designated by "max." or "min." even though it appears to be a shelf or the inner edge of a shelf, respectively, and that in the text the numbers assigned to the rings correspond to those for maxima in the tables.) The visual intensity estimates for the third and fourth rings have the ratio 4/3; this is not in accordance with the curves with C-H = 1.12 Å or more, which make the fourth ring as strong as the third. Moreover, the 1.03 curve (as well as the 1.00 curve) shows only a very poorly defined shelf in place of the fourth ring. The 1.06 and 1.09 curves reproduce in a reasonably satisfactory way all of the features observed on the photographs, including the fifth and sixth rings. The quantitative comparison of observed s values with those calculated for C-C' = 1.39 Å and C-H = 1.075 Å (the mean for the 1.06 and 1.09 curves) given in Table I, leads to C-C = 1.393 $\pm$ 0.02 Å and C-H = 1.08 $\pm$ 0.04 Å.

Wierl (4) and Pauling and Brockway (2) have previously reported C-C = 1.39 Å in benzene. Jones (5) has recently given the values C-C = 1.40 $\pm$ 0.01 Å and C-H = 1.14 $\pm$ 0.01 Å. It seems probable that the latter value is in error.

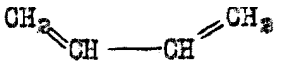
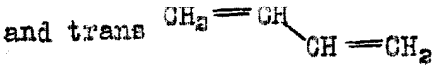
The photographs obtained for pyridine and pyrazine are so closely similar to those of benzene as to leave nodoubt that the three molecules

have nearly identical structures. The radial distribution curve for pyridine (Fig. 1) calculated from the data given in Table II has well-defined peaks at 1.38, 2.39 (= $\sqrt{3}$ · 1.38), and 2.76 (= 2·1.38) Å. The sharpness of the 2.39 peak indicates that the six meta distances in the ring are nearly equal. The calculated intensity curve for the model with C-C = 1.39 Å, C-N = 1.33 Å, and C-H = 1.08 Å is shown in Fig. 3. The comparison of $\underline{s}_o$ and $\underline{s}$ values for this model (TableIII) leads to the values C-C = 1.405 and C-N = 1.344 Å, with average ring bond distance 1.385 ± 0.01 Å. The comparison with the $\underline{s}$ values for a regular hexagonal model leads to the same value 1.385 Å for the average ring bond distance. It seems unlikely that the C-C distance differs from 1.39 Å, and the results accordingly suggest that the C-N Distance is about 1.37 ± 0.02 Å.

The principal peaks on the radial distribution curve for pyrazine (Fig. 1) lie at 1.36, 2.37 (= $\sqrt{3}$ · 1.37), and 2.76 (= 2 · 1.38) Å. Intensity curves for model I, with C-C = 1.39 Å, C-N = 1.33 Å, C-H = 1.09 Å, and all ring angles 120°, and for model II, with the same bond distances but with the ring angles at nitrogen 123° and at carbon 118°30', making the meta distances equal, are shown in Figure 3. These curves, as well as the benzene curve (representing a regular hexagonal model), agree qualitatively with the photographs. The quantitative comparison for all three leads to the value 1.366 ± 0.01 Å for the average ring bond distance. With C-C = 1.39 Å, this corresponds to C-N = 1.35 ± 0.03 Å.

Butadiene and Cyclopentadiene.- The photographs of butadiene which were obtained are diffuse and were measured only with difficulty, so that the

determination of the configuration of the molecule is less precise than usual. Thirteen rings could be seen, with the s_0 and I values given in Table IV. These correspond to the radial distribution curve shown, in Figure 1, with peaks at 1.07, 1.38, 2.12, 2.48, (3.0), 3.71, (4.13), and (4.65) Å, the values in parentheses being unreliable. The carbon-carbon peak at 1.38 Å represents the mean of the values for the two double bonds and the intervening single bond, which has some double bond character. If the double bonds have the normal distance 1.34 Å found in ethylene and allene (1b; also Eugene Eyster, Dissertation, C.I.T., 1938) the value 1.46 Å is indicated for the conjugated single bond. The peak at 2.48 Å corresponds to the 1,3 and 2,4 carbon-carbon distances; this value leads about 125° for the C=C-C bond angle.

Of the two configurations cis  and trans  for the molecule, the latter is indicated by the presence of a large radial distribution peak not at 3.0 Å, as calculated for the cis model, but at 3.71 Å, which is just the value calculated for the trans model.

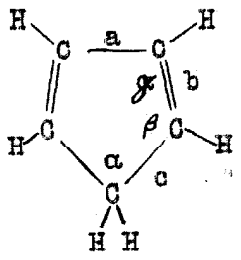
The intensity curves I, II, III, and IV of Figure 4 are calculated for trans models with C-H = 1.06 Å and with the H-C=C angles given the tetrahedral value 125°16'. The models have the following additional parameters:

Model	C=C	C-C	Angle C=C-C
I	1.38 Å	1.38	126°10'
II	1.36	1.44	124°40'
III	1.34	1.48	123°5'
IV	1.32	1.54	120°10'

The average of the carbon-carbon bond distances has for each model the radial distribution value 1.38 Å and the angle C=C-C is such as to give the radial distribution value 2.48 Å for the 1,3 and 2,4 distances. The agreement with the photographs is reasonably satisfactory for curves II and III, but not for I and IV. For example, the relative intensities of rings 9 and 10 are not well reproduced by curve IV, and those of rings 11 and 12 are not given satisfactorily by either curve I or curve IV. The appearance of an extra ring, 6, on the photographs may be due to an inflection in curves II and III in this region; otherwise the agreement is excellent. Curves were also calculated for about ten models differing slightly from these, with no improvement.

It is concluded that the butadiene molecule has the configuration described by the following parameter values: C-H = 1.07 Å, C=C = 1.35 ± 0.02 Å, C-C = 1.46 ± 0.03 Å, angle C=C-C = 124° ± 2°. The configuration is probably trans, but a small fraction of the molecules may be present in the cis form.

Data for the excellent photographs of cyclopentadiene, showing nine rings, are given in Table V. The corresponding radial distribution curve (Fig. 1) has peaks at 1.06, 1.42, and 2.33 Å. The number of parameters determining the structure of this molecule is so great as to prevent their independent evaluation. The molecule,

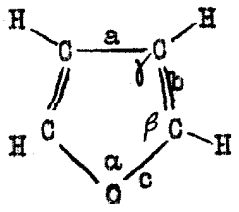


without doubt has the single bond distance c equal to about 1.53 Å.

as in the cycloparaffins, and the double bond distance b equal to about 1.35 Å, as in butadiene. The conjugated single bond c is probably about 1.46 Å. Various models with approximately these values for the carbon-carbon distances, with the ring angles close to 108°, and with C-H = 1.09 Å were found to give calculated curves agreeing closely with the photographs in appearance and leading on quantitative comparison to the value 1.445 ± 0.01 Å for the average ring bond distance. This is exactly the average of the values suggested above. Assuming the average of the values of b and c to be reliable, we write C-C = 1.46 ± 0.04 Å as the value determined for the conjugated double bond. The curve shown in Figure 4, with s values given in Table V, is that for the model with the values of a , b , and c given above and the angles $\alpha = 101° \pm 2°$, $\beta = 109° \pm 2°$, $\gamma = 110° \pm 2°$. The agreement with the photographs is less satisfactory for other models.

Furan, Pyrrole, and Thiophene.- The data for furan (Table VI) lead to a radial distribution curve (Fig. 1) with principal peaks at 1.39 and 2.26 Å, indicating an average ring bond distance of 1.39 Å. This is substantiated by comparison with calculated intensity curves, which gives 1.395 ± 0.01 Å for this average.

If the assumption be made that in the molecule



the values of a and b be 1.46 and 1.35 Å respectively, as in butadiene

and cyclopentadiene, it is found that there is excellent qualitative and quantitative agreement between the photographs and models with c = 1.41 $\pm$ 0.02 Å, α = 104° $\pm$ 3°, β=110° $\pm$ 2°, and γ = 108° $\pm$ 3°, the value C-H = 1.09 Å being assumed. The quantitative agreement is impaired by change in c, and the qualitative agreement by change in α (with corresponding changes in β and γ). It is interesting to note that the value of c is only slightly less than the single bond C—C value, 1.43 Å.

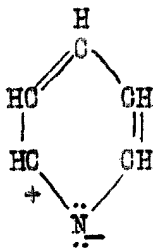
The photographs for pyrrole are closely similar to those for furan, and lead to the same value, 1.395 $\pm$ 0.01 Å, for the average ring bond distance. With a = 1.46 and b = 1.35 Å, this leads to c = 1.41 Å, which is 0.06 Å less than the C-N single bond value. In accordance with the discussion given below, it seems probable that this decrease, indicating double bond character, is accompanied by further decrease in a, and that the values in the molecule are a = 1.44 Å (assumed), b = 1.35 Å (assumed), and c = 1.42 $\pm$ 0.02 Å. The angles have the values α = 106° $\pm$ 3°, β = 109° $\pm$ 3°, and γ = 108° $\pm$ 3°. The agreement of the photographs with this model (with C-H = 1.09 Å assumed) is shown in Figure 5 and Table VII.

The data from the photographs of thiophene, showing ten rings (Table VIII), lead to the radial distribution curve of Figure 1, with peaks at 1.40 Å (C-C), 1.74 Å (S-C), 2.29 Å (C-C, meta), and 2.55 Å (S-β C). If the values C-C = 1.46 and C=C = 1.35 Å be assumed in this molecule, as in pyrrole, the radial distribution curve leads to the configuration with S-C = 1.74 $\pm$ 0.03 Å, angle C-S-C = 91° $\pm$ 3°,

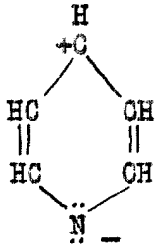
angle S-C=C = 112° ± 3°, and angle C=C-C = 113° ± 3°. The curve calculated for this model (with C-H = 1.09 Å) agrees very well with the photographs, as shown in Figure 5 and Table VIII.

Discussion

The values found for the C-N Kekule-type bond distance in pyridine and pyrazine, 1.37 ± 0.02 Å and 1.35 ± 0.01 Å, respectively, are somewhat larger than the expected value 1.33 Å for fifty percent double bond character. It is possible that this is the result of the electronegativity of the nitrogen atom, leading to significant resonance with ionic structures such as



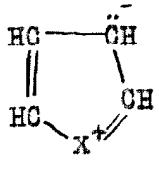
and



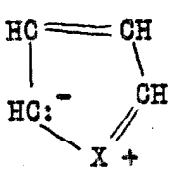
.

The value 1.46 Å for the single bond between conjugated double bonds in butadiene and cyclopentadiene corresponds to about 18% double bond character. This value agrees well with that found for a single bond between a double bond and a benzene ring in stilbene, 1.44 Å (6), in view of the fact that quantum-mechanical calculations show that a double bond and a benzene ring are about equal in conjugating power.

It is very interesting to note that the amounts of double bond character corresponding to the distances C-O = 1.41 Å, C-N = 1.42 Å, and C-S = 1.74 Å in furan, pyrrole, and thiophene are 4%, 12%, and 14%, respectively. This double bond character arises from the contribution of structures of the type



and



.

It is understandable that the very electronegative atom oxygen should not like to assume the positive charge accompanying such structures, and that in consequence resonance of this type is of little significance for furane. On the other hand, this resonance is important for pyrrole and thiophene; structures of these types make a total contribution of about 24 to 28% for these molecules.

References

(1) a. L. Pauling, L. O. Brockway, and J. Y. Beach, J.A.C.S., 57, 2705 (1935); b. L. Pauling and L. O. Brockway, ibid., 59, 1223 (1937).

(2) L. Pauling and L. O. Brockway, J. Chem. Phys., 2, 867 (1934).

(3) L. O. Brockway, Rev. Modern Phys., 8, 231 (1936).

(4) R. Wierl, Ann. Physik 8, 521 (1931).

(5) P. L. F. Jones, Trans. Faraday Soc., 31, 1036 (1935).

(6) J. M. Robertson and I. Woodward, Proc. Roy. Soc., A 162, 568 (1937).

Summary

In Section I a general radial distribution function is derived from the theoretical expression for the scattering of fast electrons by gas molecules, and its application to data obtained by visual observation of diffraction photographs is discussed. A useful modification of the usual radial distribution method is presented, together with suggestions for its further improvement and some criteria for its reliability.

The electron diffraction investigation reported in Section II leads to values for the interatomic distances and bond angles of the molecules discussed there, and to information regarding their electronic structures.

Acknowledgment

I am very grateful to Professor Linus Pauling for his continual assistance and encouragement during these investigations.

The second section of this thesis has been prepared jointly by Professor Pauling and myself. To Professor L. O. Brockway I am indebted for many helpful discussions and instruction in the technique of electron diffraction; and I wish here to thank Mr. Saul Winstein who prepared the sample of cyclopentadiene, and Mr. K. J. Palmer who made the electron diffraction photographs of this substance.

I wish to call attention to the importance of Dr. Charles Degard's contributions to the first section of this thesis, and to express my appreciation for his interest in this application of his work.

Finally, I take pleasure in thanking Dr. Sidney Weinbaum, Mr. E. H. Eyster, Mr. A. J. Stosick, and Mr. H. A. Levy for their further special assistance.

PROPOSITIONS

1. The radial distribution method[1,2] of interpretation of electron diffraction photographs of gas molecules can be modified in such a way as to increase its usefulness. The modification involves only a change in the treatment of the coefficient of the radial distribution function, and does not appreciably increase the laboriousness of the calculation. It is desirable to use data from <u>both</u> the observed "maxima" and "minima."

2. As the inevitable extension of the electron diffraction method to the study of more and more complicated molecules proceeds, the computational labor involved in the interpretation of the data by the more important visual method[3] becomes increasingly burdensome. It is therefore appropriate to consider the following means of decreasing this labor:

a) Calculation of the function

$$I'(s) = \sum_{ij}{}' \frac{C_{ij}}{\ell_{ij}} \sin \ell_{ij} s$$

instead of $I(s) = \sum_{ij}{}' C_{ij} \sin \frac{\ell_{ij} s}{\ell_{ij} s}$

b) Use of a rapid "Fourier synthesizer."

c) Use of a plot, showing the locus of the maxima of $\xi \sin \mathscr{e}_{ij} S \xi$ (as a function of $\mathscr{e}_{ij}$ and S) as an aid in the choice of models to be considered.

d) In the case of "Free Rotation" models, the use of a two or four center Gaussian representation of the distribution of scattering matter, together with the "temperature effect" formulae.

3. The theoretical treatment of James[3] of the effect of molecular vibration on the diffraction of x-rays (or fast electrons) by gas molecules is unnecessarily complicated, and may easily be simplified.

4. Contrary to previous statements[4], the "temperature effect" is sometimes important in the study of molecular structure by the electron diffraction method. There are, roughly, four cases to be considered and examples of them have been found[5]:

1. There exist important interatomic separations which do not have reasonable definite values.
2. The molecule has a reasonably definite configuration, but is large and complicated.
3. Certain "force constants" of the molecular potential function are unusually small.

4. The molecule contains many hydrogen atoms together with heavier atoms.

5. The discussion[6] of ionic radii given by Zachariasen is defective, and consequently differences between his set of ionic radii and the earlier set[7] by Pauling should not be regarded as being very significant.

6. The apparently normal heat of hydrogenation of hydrindene[8] need not be regarded as contradictory to the thesis of Sidgwick and Springall[9] that in this compound the benzene system is considerably perturbed by the Mills Nixon effect.

7. Comparison of the observed Raman frequencies[10] of liquid phosphorus pentachloride with the preliminary results of a normal coordinate treatment for a trigonal bipyramidal model indicates that for PCl_5 a simple valence force potential function is unsuitable.

8. In the gas phase molecules of phosphorus pentachloride have a symmetrical trigonal bipyramidal structure in which the equatorial bonds are about 5% shorter than the axial bonds. The first of these statements, but not the second, is in agreement with the predictions of the directed valence theory of Pauling.

9. The "contribution" of "excited" Lewis structures to the ground state of simple heterocyclic molecules can be estimated from observed interatomic distances and may be satisfactorily correlated with the properties of the hetero atoms involved.

10. The effect on interatomic distance of one or three electron bond formation in resonating systems of the type (—— ⇌ ···), etc., is approximately half that of two electron bond formation in the analogous systems (—— ⇌ ——) .., etc.

11. Simple considerations based on the quantum mechanical treatment of molecules lead to the prediction that carbon dioxides (C_nO_2) containing an <u>odd</u> number of carbon atoms should be more stable than those containing an even number. This is in agreement, at least, with the fact that CO_2, C_3O_2, and C_5O_2[11] have been prepared while C_2O_2, C_4O_2, etc. have not.

Similarly it may be expected that the tetraphenyl cumulenes $\{\varphi_2C{=}(C)_n{=}C\varphi_2\}$ and similar molecules should fit according to their chemical properties into two series consisting of the members with n even and n odd. The <u>even</u> members should be the more stable.

13_1. Even isotopic diatomic molecules, e.g., HD, should exhibit pure-rotation and rotation-vibration absorption spectra. The bands would not be very intense.

13_2. Very meager, but nevertheless impressive, experimental evidence suggests that the reactivity of mixed phosphorus trihalides containing fluorine depends on the number of fluorine atoms in the molecule in a way similar to that found for the halogenated methanes[12].

13_3. The advantages afforded teaching fellow guests by the Athenaeum have been exaggerated in various official published statements concerning it.

1. Linus Pauling and L. O. Brockway, J. A. C. S. 57, 2684 (1935)
2. L. O. Brockway, Rev. Mod. Phys. 8, 250, 256 (1936)
3. R. W. James, Physik Zeit 33, 737 (1932)
4. L. O. Brockway, Rev. Mod. Phys. 8, 239 (1936)
5. 1) $\alpha\omega$ dihalogenoalkanes (Ketelaar and Palmer, J. A. C. S. (59, 2470 (1937)
   2) $P_3N_3Cl_6$ (Sohomaker)
   3) Ge_3H_8 (Pauling, Laubengayer, and Hoard)
   Tetranitromethane (Stosick)
   4) Ethylene Oxide (Lévy)
6. Z. Krist, 80, 137 (1931)
7. J. A. C. S. 49, 765 (1927)
8. Dolliver, Gresham, Kistiakowsky and Vaughan, J. A. C. S. 59, 831 (1937)
9. Chemistry and Industry 55, 476 (1936)
10. Moreu, Magat and Wetroff, Compt. Rend. 205, 276 (1937)
11. Klemec and Wagner, Ber. P 1886 (1937)
12. See for example, L. O. Brockway, J. Phys. Chem. 41, 185 (1937)

ILLUSTRATIONS AND TABLES

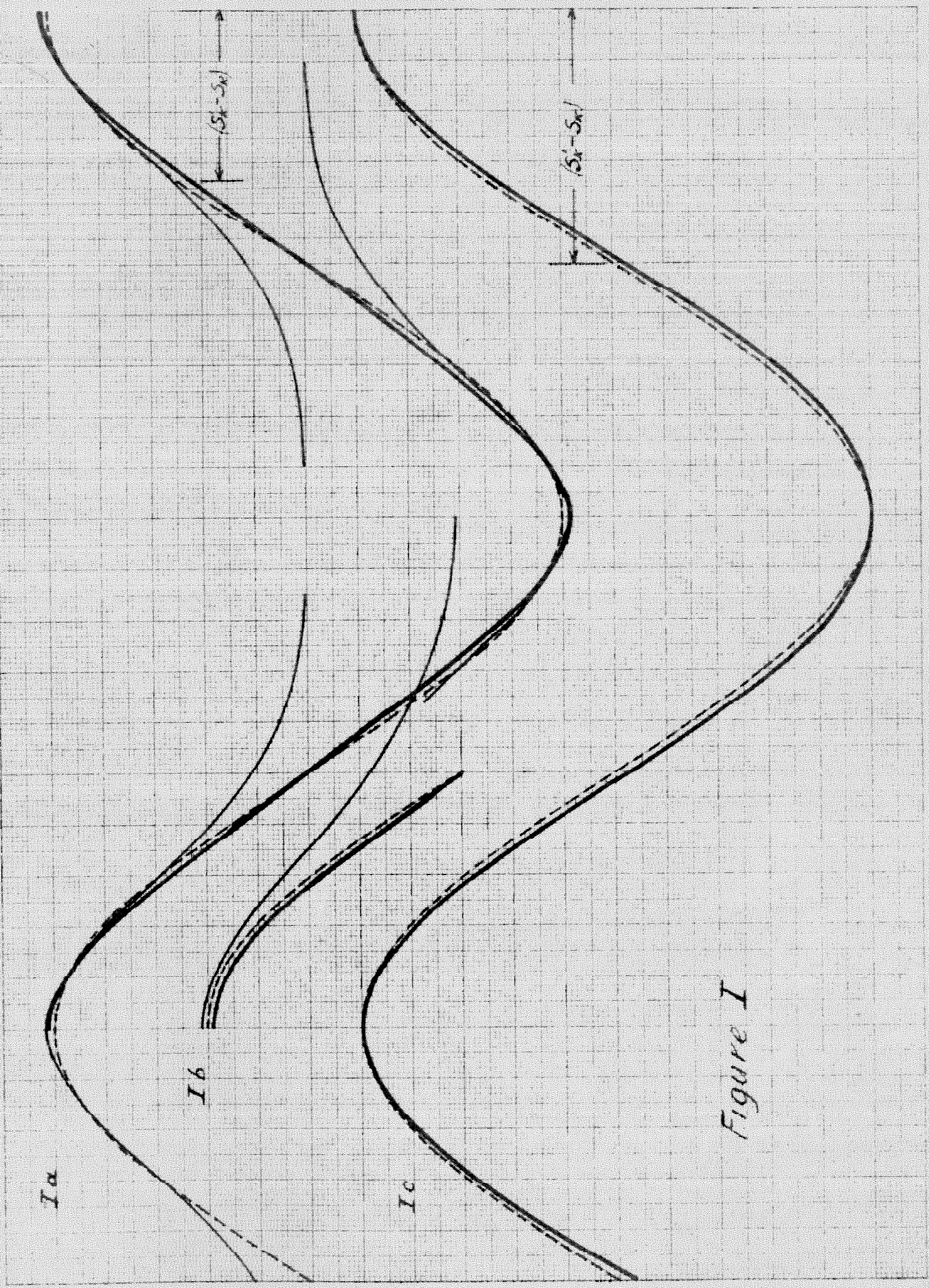


Figure I

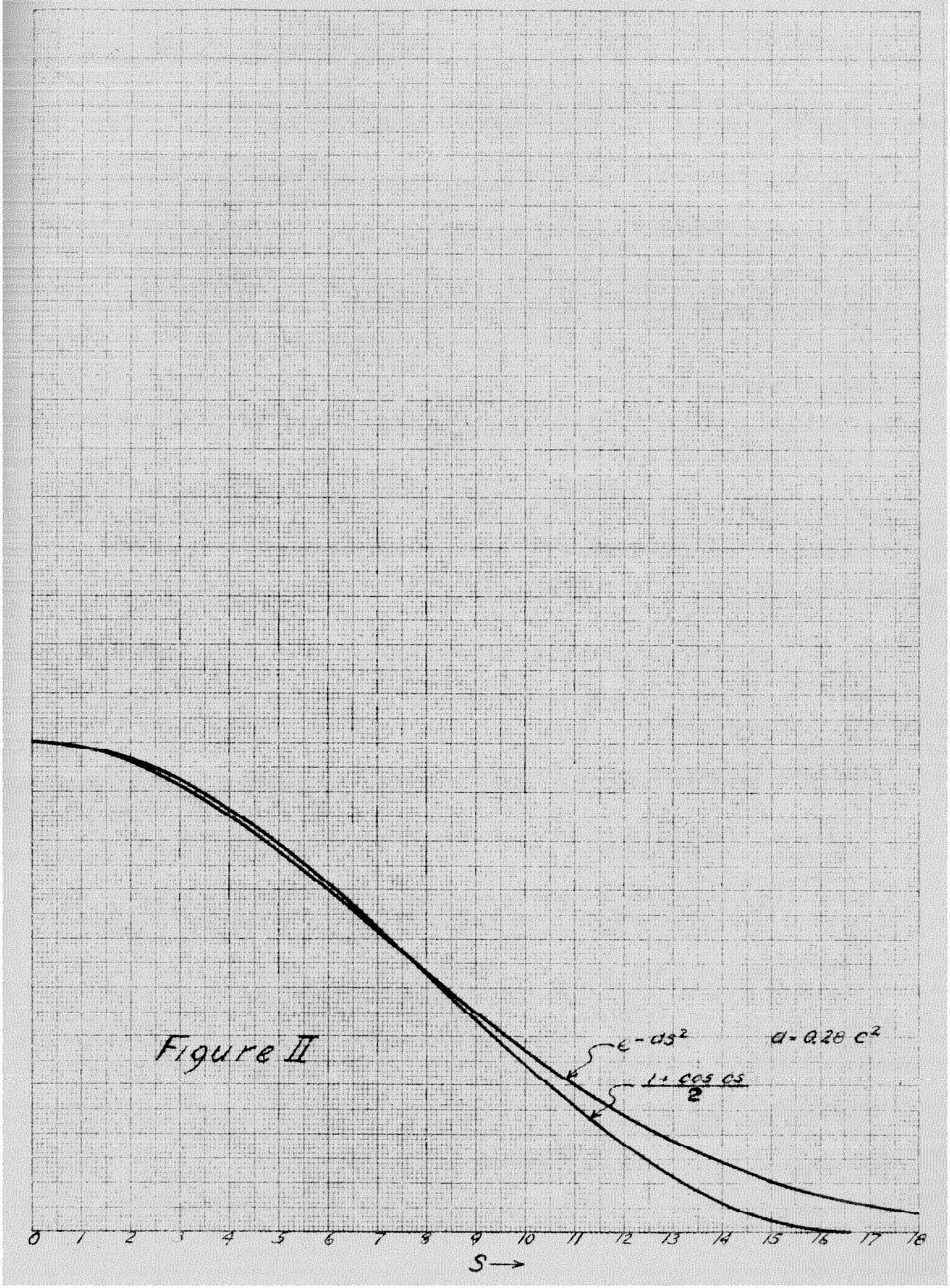

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
S →
$\epsilon - as^2$
$a = 0.28\ c^2$
$\frac{1+\cos cs}{2}$

Figure II

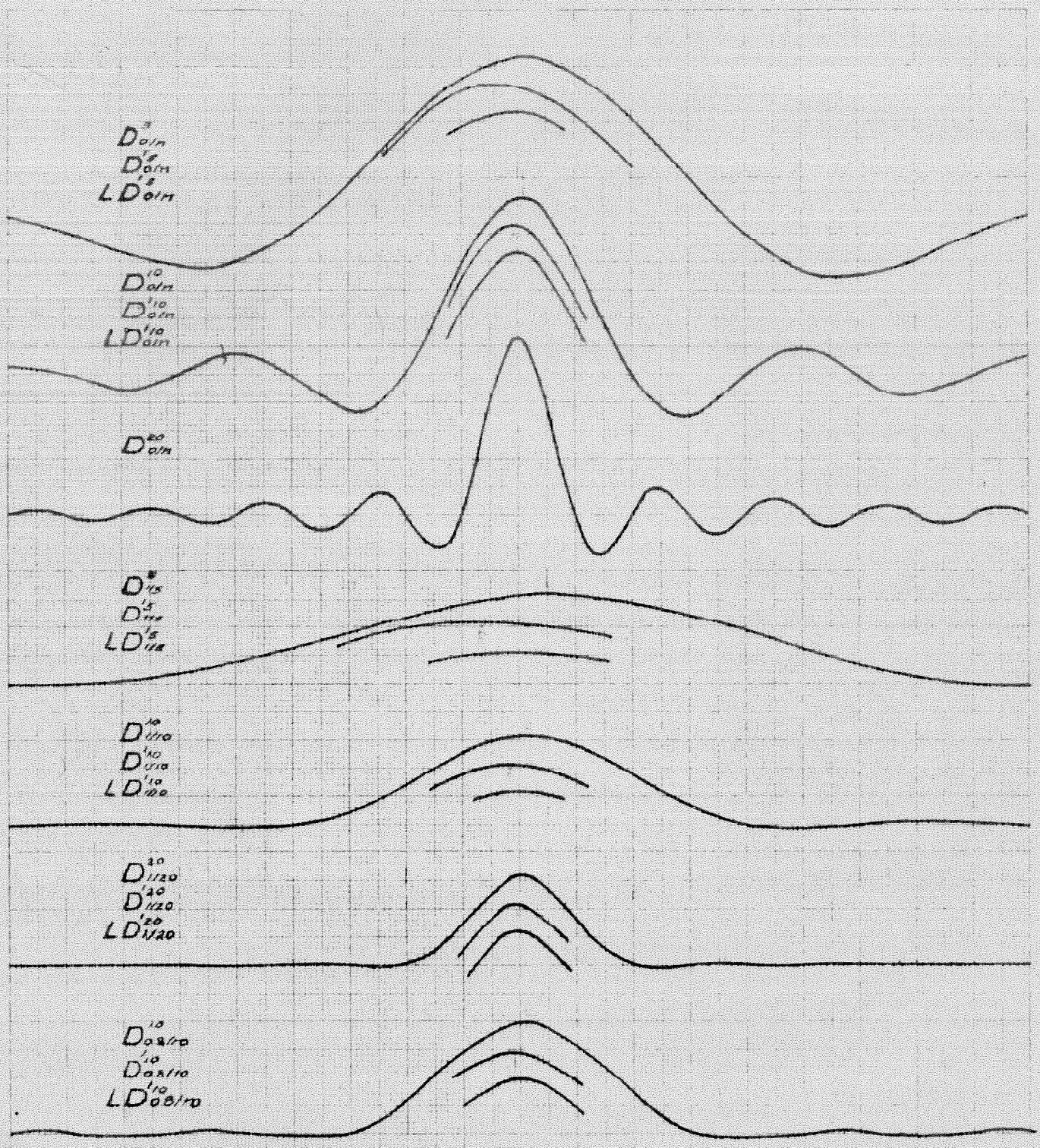


Figure III

0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4

L →

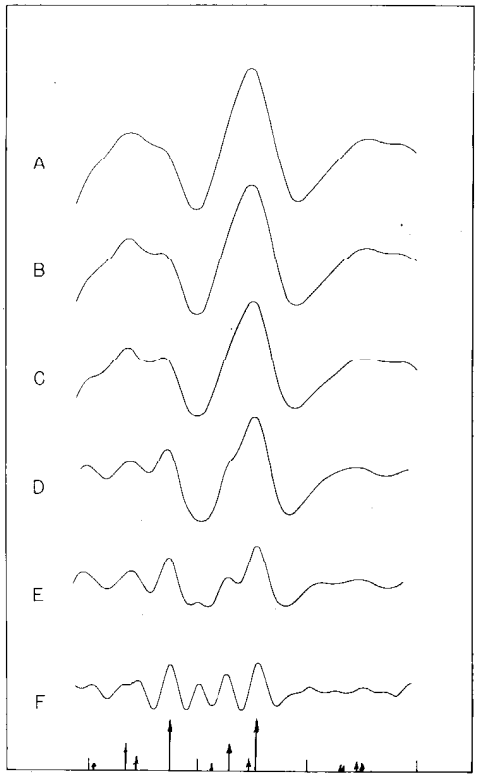


Figure IV. Radial Distribution Curves for Thiophene

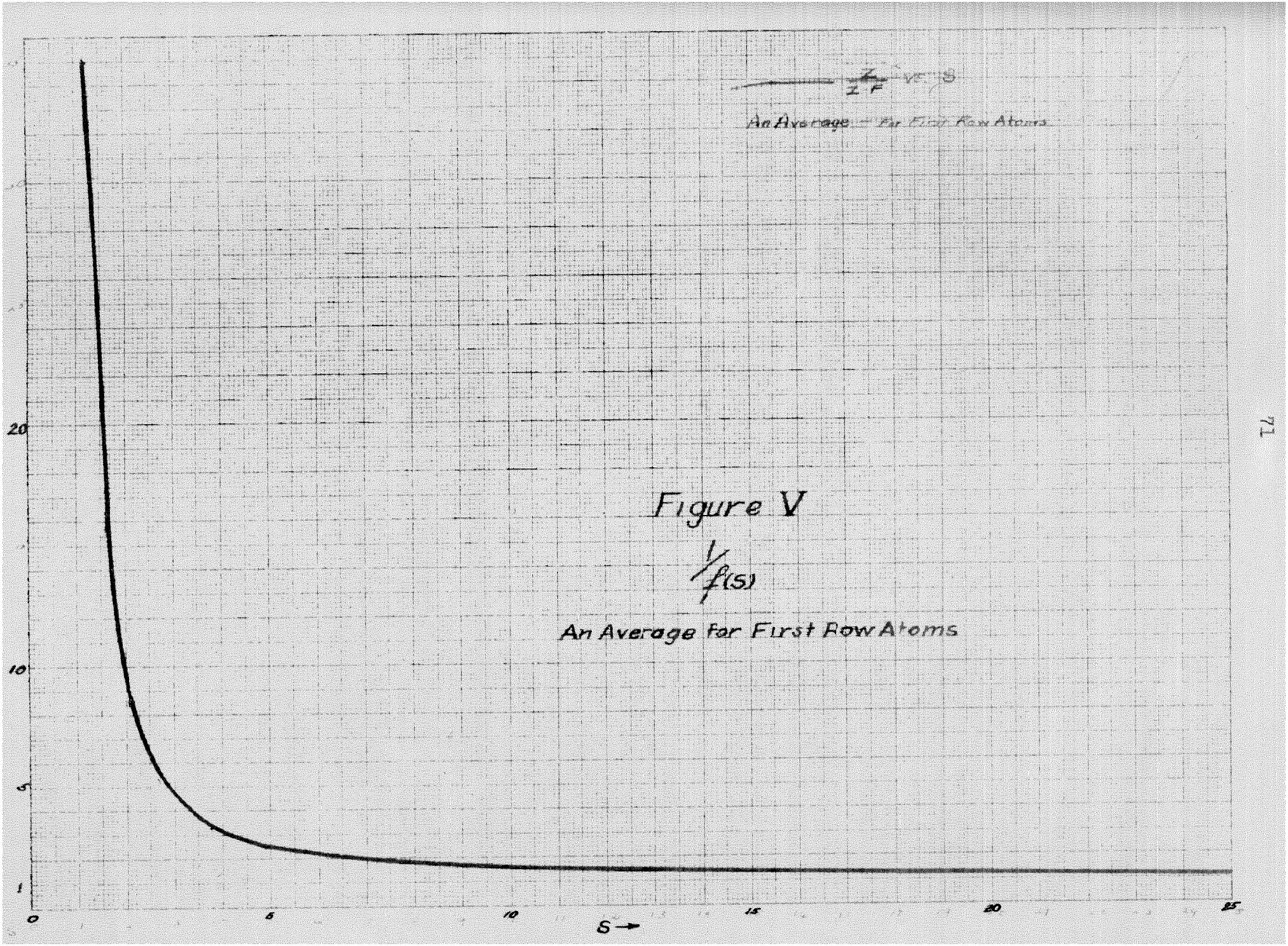


Figure V

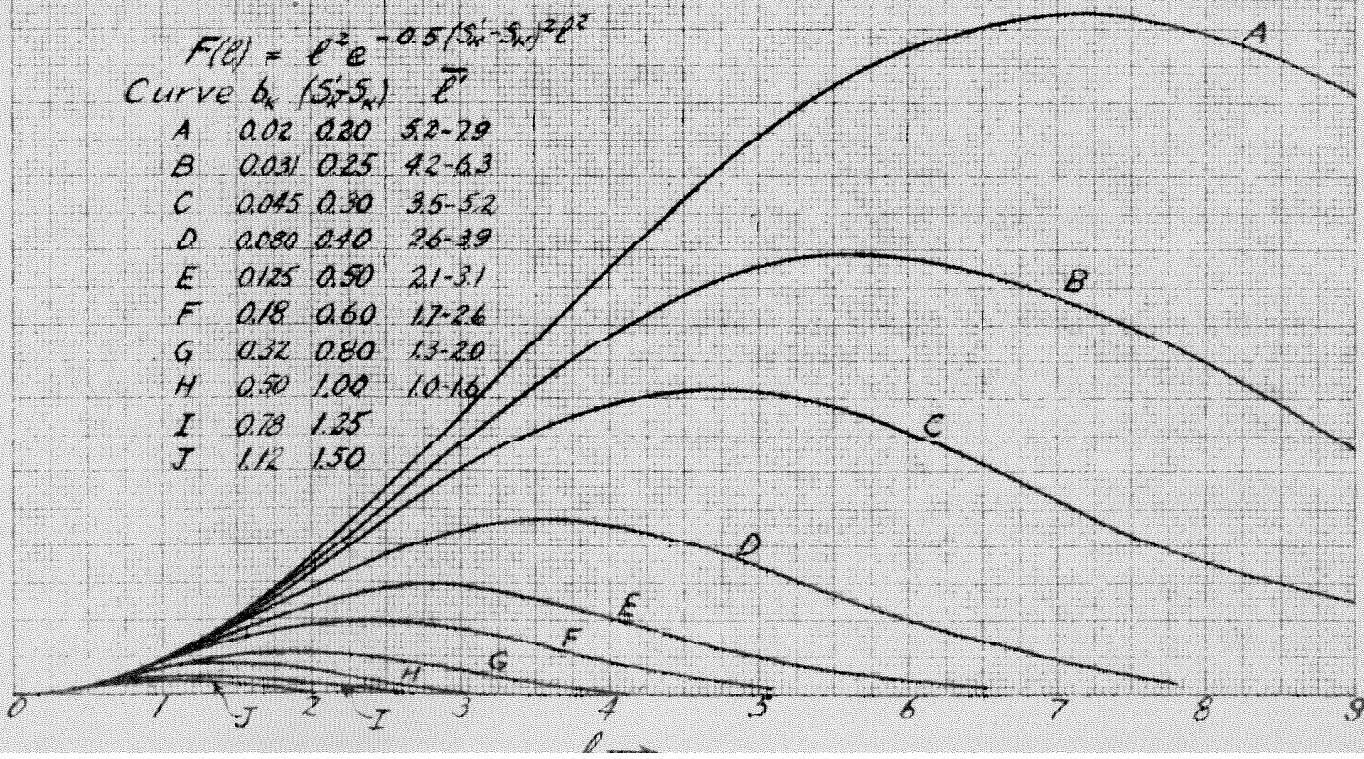
$\frac{1}{f(s)}$

An Average for First Row Atoms

Figure VI

$F(\ell) = \ell^2 e^{-0.5(S_k - S_{k-1})^2 \ell^2}$

Curve	b_k	$(S_k - S_{k-1})$	$\bar{\ell}$
A	0.02	0.20	5.2-7.9
B	0.031	0.25	4.2-6.3
C	0.045	0.30	3.5-5.2
D	0.080	0.40	2.6-3.9
E	0.125	0.50	2.1-3.1
F	0.18	0.60	1.7-2.6
G	0.32	0.80	1.3-2.0
H	0.50	1.00	1.0-1.6
I	0.78	1.25	
J	1.12	1.50	



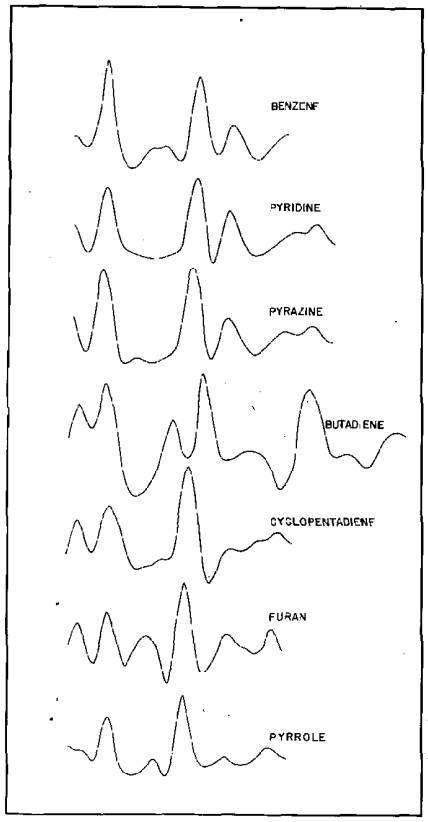


Figure I

Radial Distribution Curves

Coefficients:

Benzene, Pyridine, Pyrazine, and Butadiene; $C_K = I_K s_K^2 e^{-as_K^2}$. Both maxima and minima.

Cyclopentadiene; $C_K = I_K s_K^3 e^{-as_K^2}$. Maxima only.

Furan and Pyrrole; $C_K = I_K s_K^4 e^{-as_K^2}$. Maxima only.

Thiophene; See figure IV, section I. $C_K = I_n s_K^2$. (No artificial temperature factor)

Table I

Benzene

Max.	Min.	I	C_K	s_o	s^a	s/s_o
1		4	2	3.20	2.89	(0.904)[b]
	2	-10	-7	4.30	4.19	(0.974)
2		10	14	5.71	5.75	1.006*
	3	-1	-2	7.60	7.57	0.996
3		4	11	9.48	9.89	(1.043)
	4	-1	-3	10.26	10.92	(1.063)
4		3	11	11.20	11.15	(0.991)
	5	-7	-27	12.46	12.50	1.003*
5		4	19	13.63	13.89	1.020
	6	-2	-9	14.85	---[c]	---
6		2	10	15.98	15.80	0.989
	7	-5	-23	17.42	17.40	0.999*
7		3	15	18.61	18.84	1.012*
	8	-½	-2	20.18	20.29	1.005
8		1½	7	21.30	21.17	0.994
	9	-2	-8	22.70	22.20	0.979
9		2	8	23.80	24.00	1.008
10		1	3	(27.0)	---[c]	---

Average 1.002

Average Deviation .008

(C-C = 1.39 x 1.002 = 1.393 Å)
(C-H = 1.075 x ----- = 1.077 Å)
C-C = 139 ± 0.02Å
C-H = 1.08 ± 0.04Å

[a] Calculated for plane, regular hexagonal model with C-H = 1.075 Å and C-C = 1.39 Å or (See text for description of model)

[b] Less reliable values, not included in taking the average, are shown in parentheses.

[c] The theoretical curve does not show a measurable feature at this point.

* In finding the average value of s/s_o, the starred values are given double wt. for we believe them to be more reliable than the others.

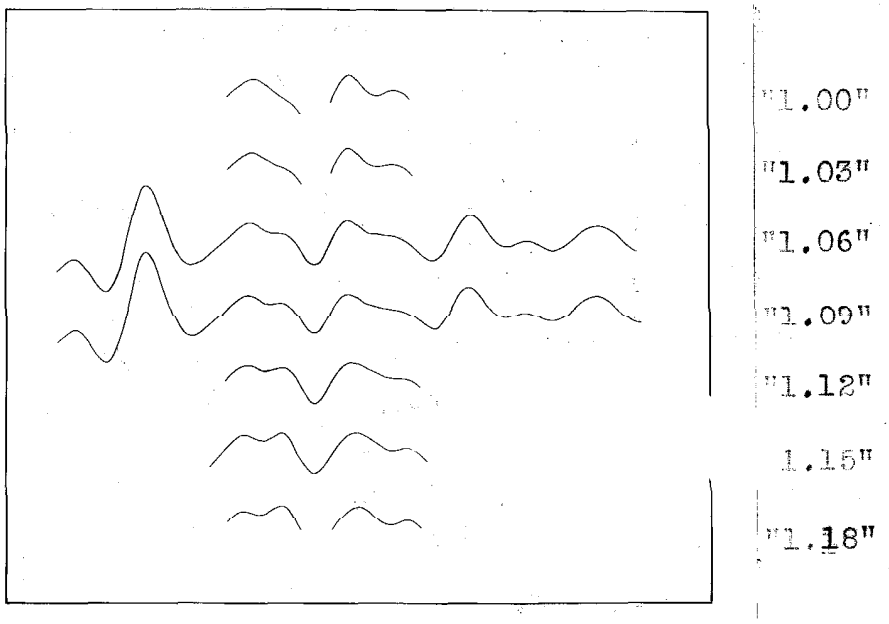


Figure II

Theoretical Intensity Curves for Benzene

TABLE II. Pyridine

				Model I[a]			Benzene "1.09, ""1.59"	
Max.	Min.	I	C	s_o	s	s/s_o	s	s/s_o
1		4	2	3.19	2.99	(0.947)	2.89	(0.906)
	2	-10	-7	4.37	4.16	(0.952)	4.19	(0.959)
2		10	14	5.69	5.83	1.025*	5.75	1.011*
	3	-1	-2	7.56	7.68	1.027	7.57	0.999
3		4	11	9.14	9.94	(1.087)	9.89	(1.082)
	4	-1	-3	10.21	10.74	(1.053)	10.92	(1.070)
4		3	11	11.24	11.30	(1.004)	11.15	(0.992)
	5	-7	-27	12.73	12.75	1.002*	12.50	0.982*
5		4	19	13.88	14.23	1.025	13.89	1.001
	6	-2	-9	15.12	15.85	(1.048)	---	---
6		2	10	16.20	16.31	1.007	15.80	0.975
	7	-5	-23	17.60	17.66	1.003*	17.40	0.988*
7		3	15	18.86	19.10	1.013*	18.84	0.999*
	8	-½	-2	20.33	20.60	1.013	20.29	0.997
8		1½	7	21.44	21.40	0.997	21.17	0.987
	9	-2	-8	22.65	22.47	0.992	22.20	0.980
9		2	8	23.82	24.27	1.018	24.00	1.007
10		1	3	28.0	28.18	1.007	---	
			Average			1.011		0.994
			Average Deviation			0.009		0.010

[a] Models described in text.

Table III

Pyrazine

					Model I[a]		Model III		Benzene	"1.09"1.39"
Max.	Min.	I	c_A	s_o	s	s/s_o	s	s/s_o		
1		4	2	3.09	3.02	(.977)	3.06	(.990)	2.90	(.938)
	2	-10	-10	4.41	4.23	(.959)	4.20	(.952)	4.19	(.950)
2		10	16	5.89	5.87	.996*	5.89	1.000*	5.75	.976*
	3	-2	-5	7.59	7.77	1.023	7.71	1.016	7.60	1.002
3		4	12	9.08	10.23	(1.127)	10.24	(1.128)	9.87	(1.087)
	4	$-\frac{1}{8}$	-2	10.20	10.70	(1.049)	10.62	(1.042)	10.90	(1.068)
4		4	16	11.27	11.37	(1.008)	11.35	(1.007)	11.20	(.994)
	5	-7	-31	12.76	12.90	1.012*	12.89	1.011*	12.50	.980*
5		4	19	14.07	14.43	1.026	14.45	1.027	13.93	.991
	6	-1	-4	15.21	16.02	(1.053)	15.92	(1.047)	---[c]	---
Shelf 6		2	10	16.20	16.50	1.018	16.55	1.022	15.80	.975
	7	-5	-25	17.73	17.90	1.010*	17.95	1.007*	17.40	.982*
7		3	14	19.16	19.40	1.013*	19.41	1.013*	18.85	.984*
	8	-1	-4	20.37	20.85	1.024	20.90	1.026	20.30	.997
8		1	4	21.51	21.60	1.004	21.70	1.008	21.13	.983
	9	-2	-7	22.97	22.70	.988	22.78	.992	22.20	.967
9		2	6	24.04	24.55	1.021	24.65	1.025	24.00	.998
	Average					1.012		1.012		.984
	Average deviation					.008		.008		.007

[a] Models described in text.

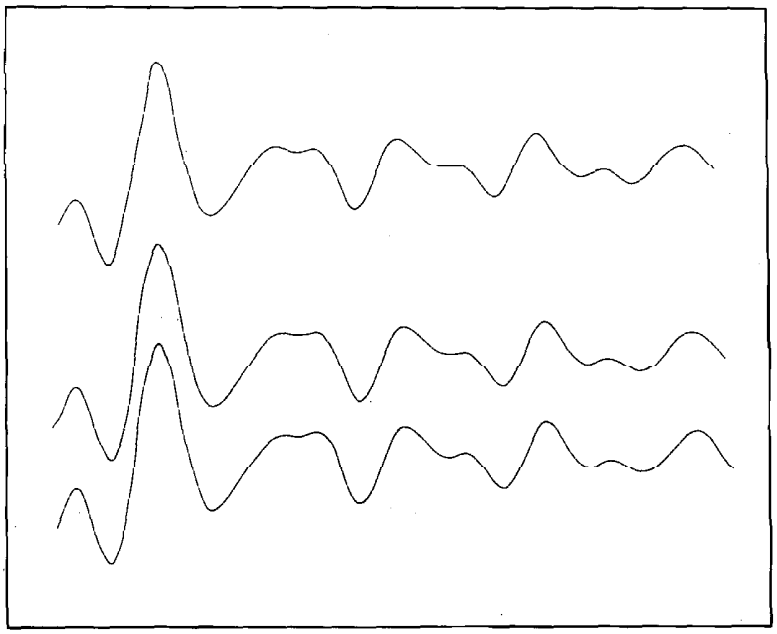


Figure III

Theoretical Intensity Curves
for Pyridine and Pyrazine

TABLE IV

Extadiens

Max.	Min.	I	Δ_K	a_0	Model II a	Model II a/a_0	Model III a	Model III a/a_0
	1	-3	0	2.37	--	(-----)	---	(---)[a]
1		5	2	3.35	--	(-----)	--	(-----)
	2	-10	-7	4.47	4.01	(0.897)	4.10	(0.918)
2		15	14	5.62	5.92	(1.053)	5.94	(1.057)
	3	-1	-1	6.24	---	(-----)	---	(-----)
3		7	11	6.98	---	(-----)	---	(-----)
	4	-2	-4	7.74	7.90	1.020	7.95	1.026
4		4	8	8.79	9.00	1.023	9.05	1.029
	5	-1	-2	9.68	9.30	0.961	9.40	0.971
5		6	16	10.56	1.062	1.006*	10.40	0.985*
	6	-4	-11	11.67	11.55	0.990*	11.60	0.994*
6		3	10	12.74	---	(-----)	---	(-----)
	7	-1	-3	13.33	---	(-----)	---	(-----)
7		3	9	13.82	---	(-----)	---	(-----)
	8	-2	-6	14.60	---	(-----)	---	(-----)
8		3	10	15.58	---	(-----)	---	(-----)
	9	-4	-13	16.90	16.92	1.001	16.93	1.002
9		3	10	18.37	18.68	1.017	18.69	1.017
	10	-1	-3	19.72	19.94	1.011	19.80	1.004
10		1	3	20.79	20.40	0.981	20.72	0.947
	11	-2	-5	22.13	22.00	0.994	22.35	1.010
11		2	4	23.88	23.97	1.004	24.12	1.010
	12	-1	-2	24.87	---	(-----)	---	(-----)
12		1	2	26.26	---	(-----)	---	(-----)
	13	-1	-1	27.10	---	(-----)	---	(-----)
13		1	1	28.20	---	(-----)	---	(-----)
			Average			1.001		1.003
			Average Deviation			0.012		0.011

[a] The theoretical curve does not have a well defined max. or min. corresponding to the measured ring.

TABLE V

Cyclopentadiene

Max.	Min.	I	C_k	s_o	Model I[a] S	Model I[a] s/s_o
1		28	1	3.66	—	—
	2			4.71	4.45	(0.946)
2		70	14	6.13	6.00	(0.979)
	3			7.54	7.46	0.990
3		28	15	8.98	9.01	1.003
	4			10.37	10.41	1.004
4		12	11	11.54	11.53	0.999
	5			12.78	12.68	0.992
5		16	24	14.09	14.14	1.003
	6			15.68	15.80	1.007
6		7	12	16.90	17.07	1.009
	7			18.16	18.15	0.999
7		7	13	19.29	19.40	1.006
	8			20.99	21.40	1.019
8		3	5	22.28	22.45	1.007
	9			23.62	23.43	0.992
9		3	5	24.87	24.86	0.999
			Average			1.002
			Average Deviation			0.005
			Average Ring Band			1.443 Å

[a] Model for which a = 1.44 Å
b = 1.34 Å
c = 1.54 Å
= 1.01°

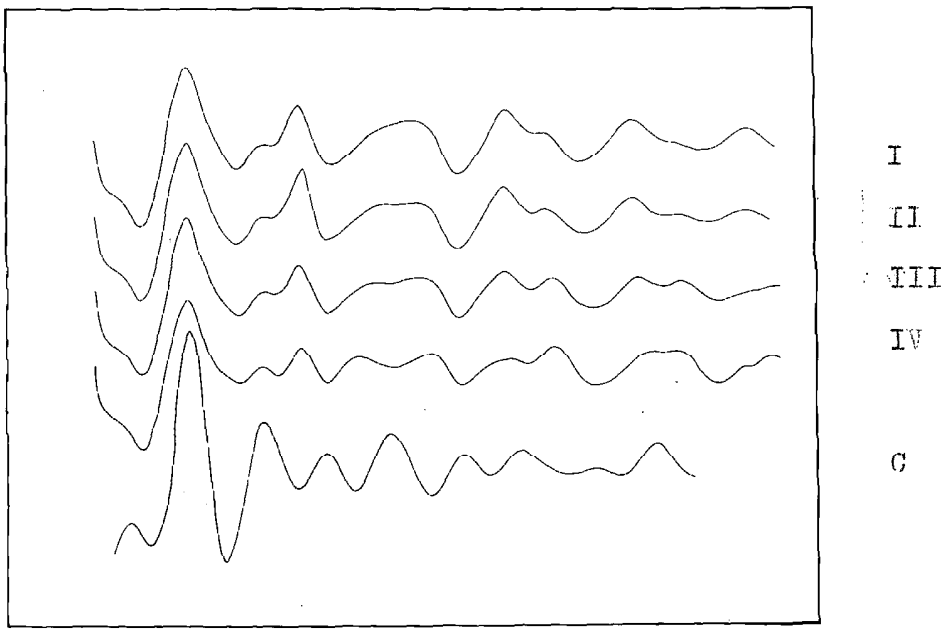


Figure IV

Theoretical Intensity Curves
for Butadiene (I, II, III, IV),
and Cyclopentadiene (C).

TABLE VI

Furan

Max.	Min.	I	C_k	s_o
	1			2.42
1		11	0	3.61
	2			4.78
2		20	4	6.15
	3			7.76
3		12	7	9.13
	4			10.60
4		6	15	11.73
	5			13.16
5		10	28	14.44
	6			16.23
6		3	13	17.65
	7			18.89
7		3	15	19.88
	8			21.55
8		1½	11	23.26
	9			24.26
9		1½	5	25.46

TABLE VII

Pyrrole

Max.	Min.	I	C_k	s_o	Model I[a] s	s/s_o
1		10	0	3.61	---	(------)
	2			4.85	---	(------)
2		30	8	6.26	6.09	0.972
	3			7.79	7.71	0.989
3		12	7	9.36	9.49	1.014
	4			10.62	10.83	1.019
4		5	7	11.79	11.81	1.001
	5			13.12	13.04	0.994
5		10	22	14.53	14.67	1.009
	6			16.18	16.47	1.018
6		4	13	17.54	17.85	1.017
	7			18.72	18.77	1.003
7		4	15	19.93	19.95	1.001
	8			21.60	21.79	1.009
8		4	17	22.98	23.38	1.017
	9			24.57	24.65	1.003
9		1	4	25.82	25.70	0.995
	10			27.14	27.15	1.000
10		2	7	28.07	28.60	(------)
	11			29.34	30.19	(------)
11		1	3	30.81	31.55	(------)

Average	1.004
Average Deviation	0.010
Average Ring Band	1.395 Å

[a] Model: a = 1.47 Å
b = 1.34 Å
c = 1.40
d = 105°

TABLE VIII

Thiophene

Max.	Min.	I	C_k	s_o	Model I[a] s	s/s_o
1		2	0	3.37	3.53	(1.047)[b]
	2			4.42	4.50	(1.019)
2		10	5	5.64	5.52	(0.978)
	3			6.90	6.80	0.986
3		8	9	8.17	8.05	0.985
	4			9.53	9.50	0.997
4		3	5	10.71	10.92	(1.019)
	5			11.70	12.08	(1.032)
5		1	3	12.63	12.42	(0.983)
	6			13.80	13.78	0.999
6		3	11	15.05	15.30	1.016
	7			16.68	16.82	1.008
7		1	5	17.98	18.30	(1.017)
	8			18.88	18.85	(0.998)
8		1	6	19.71	19.60	(0.994)
	9			21.10	21.30	1.009
9		1	9	22.53	22.79	1.010
	10			23.97	24.04	1.003
10		$\frac{1}{2}$	5	25.41	25.33	0.997
			Average			1.001
			Average Deviation			0.008

[a] Model described in text

[b] Many of the features of the thiophene photographs are of types known to be generally hard to measure. Here the inner rings, and the components of the two doublets have been thrown out of the average, since measurements of both types of features are known to be unreliable. The effect of the choice of rings to be thrown out is here somewhat greater than usual and may amount to $\pm$ 0.005 in s/so.

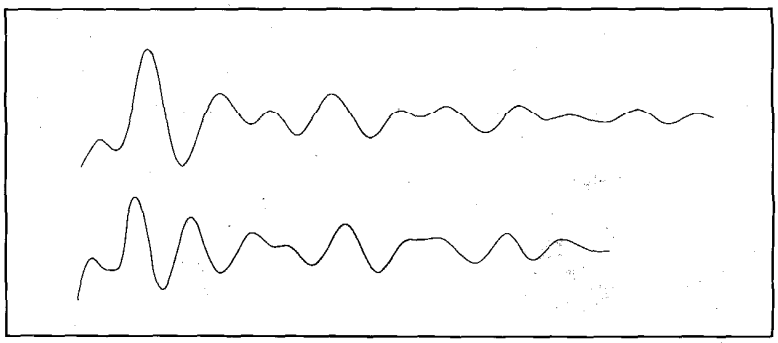


Figure V

Theoretical Intensity Curves
for Pyrrole and Thiophene